INTERNATIONAL CONFERENCE ON ENGINEERING DESIGN, ICED'09
24 - 27 AUGUST 2009, STANFORD UNIVERSITY, STANFORD, CA, USA

LOWERING BARRIERS TO DISTRIBUTED DESIGN RESEARCH COLLABORATION

David M Sirkin, Neeraj Sonalkar, Malte Jung and Larry J Leifer
Stanford University, Center for Design Research

ABSTRACT
This paper presents experiments that were undertaken at the Stanford Center for Design Research to enable remote participants to engage in weekly design research gatherings. The context of these sessions had a number of individual participants at multiple locations and a large group at Stanford where the sessions were held – an interaction we term 1-to-1-to-many. Each session varied between 3 major categories of exchange: informal socializing, formal presentations and dynamic Q&A dialog. This tested current video conferencing technology to its limits and necessitated exploratory prototyping to address issues of social and content-rich interaction between remote and local participants. We found (1) that providing a communication channel for the remote participants to talk amongst themselves helped to create a strong sense of community distinct from that of the larger group, (2) that aggregating the representations of remote participants to a single shared display diminished the sense of each individual's persona, but disaggregating them without sufficient supporting technology led to a perception of disembodiment, and (3) that these effects were magnified during informal exchanges.

Keywords: communication technologies, virtual organizations, distributed communities, informal interaction, remote collaboration.

1 NEED: SUPPORTING A DISTRIBUTED COMMUNITY
As communities of practice become increasingly global, they face the challenge of enabling social interaction over a distance. Communities of practice as proposed by Lave and Wenger [1] are formal or informal communities that revolved around a shared practice. These could be a group of designers in a firm or, as presented here, a community of design researchers in an educational institution. Within such contexts, the learning process occurs through the community engaging in social interactions rather than by isolated individuals. This is especially crucial for the newcomers in the community. The consequences of reduced social interaction could vary from a short-term loss of content exchange about the task at hand to a more long-term loss of context shared by the community.

Prior research relevant to the area of enabling distributed communities of practice includes work done from an organizational perspective in distributed knowledge management [2], learning perspective in distance learning [3] and technology perspective in distributed communication in human computer interaction [4]. In this paper, we approach the topic from a technology perspective by focusing on the design of a synchronous communication system for enabling social interactions in a distributed design research community.

2 CONTEXT: REMOTE PARTICIPANTS AT DESIGNX
DesignX is a weekly one-and-a-half hour gathering held at the Center for Design Research at Stanford University. During any week's session, one or two faculty, ten to twenty students and several guests discuss the latest design community happenings and research activity.

Over the last two years, designX has seen an increasing number of global participants. Up to five members join each week from locations including Canada, Japan, Germany and the East and West coasts of the US. Each of these remote participants is a single individual at a distinct location. This presents a 'satellite' communication setup, where several remote individuals are connected to a single

central group. Over time, the setup has grown to include direct connections among the remote participants as well, as shown in Figure 1. We call this mode of communication '*1-to-1-to-many*,' extending the variables used for analyzing organizational use of communication technology presented in Nass and Mason [5].

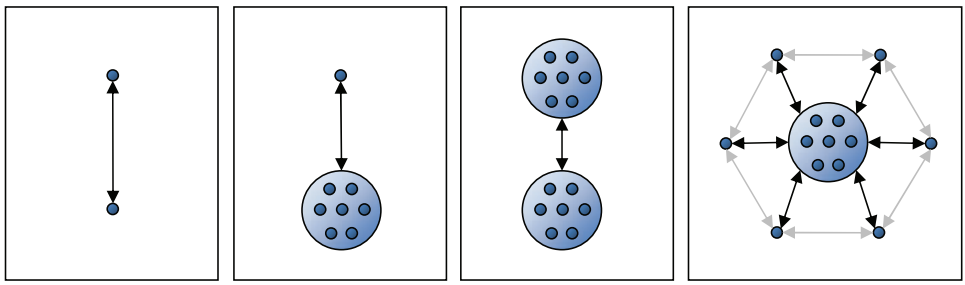

Figure 1 – The left 3 modes of communication are (from L-R) 1-to-1, 1-to-many and many-to-many. In each case, there is a single channel between the individuals or groups at each end of the connection. The mode of communication on the right is what we term 1-to-1-to-many. There is a separate channel for each individual to connect to the group, as well as a back-channel to connect to other individuals.

Another noteworthy aspect of communication at designX is its variation between three categories of social exchange. Sessions begin with informal socializing over food and drink, continue with a more formal structured presentation, and conclude with a semi-structured dynamic Q&A dialog. These transitions result in remote participants – much like their local counterparts – assuming several roles in settings that vary from close-up small-group chat, to presentation with overheads, to room-scale large-group discussion.

The increasing participation of remote members at designX sessions highlighted the insufficiency of the currently-available social communication systems to handle the breadth of usage circumstances. Depending on the setup employed, remote participants were perceived by the local group either as passive observers lurking the fringe of activity, or as visible collaborators who required significant effort to include in the conversation. From the perspective of the remote participant, the local group was barely aware of their existence. The satellite communication setup and variation in social role have direct implications for why the communication system failed, and why we chose to explore alternative approaches.

We sought a solution that could support participants in establishing common ground [6, 7] across each of the categories of social exchange. And while our goal was not explicitly to create a sense of physical or social presence [8] among members of the community, we readily discarded alternatives that broke that sense. The following section presents the 'problem-finding' approach we took to designing a system that could encourage mutual awareness, more productive interactions and a stronger sense of community.

3 EXPLORATIONS: COMMUNICATION PROTOTYPES

Our explorations comprised a series of communication prototypes, each of which attempted to address the shortcomings of the existing technology or the preceding iteration. The goal of these explorations was to develop a system which would enable local and remote participants to interact effectively while keeping the load on the technology mediators to a minimum. Following are descriptions of each prototype, along with motivations for their design and insights gleaned from their (often brief) use.

Table 1 – A summary of design research communication explorations. Each exploration consisted of a prototype to address one or more of several variables that influence effective communication. The outcome of each experiment was often mixed, with improvements in one aspect of communication coming at the cost of another. But each provided insights that could be used in the next prototype.

Exploration Prototype	*Variables Affected*	*Behavioral Insights*
High-end video conference system	Communication channel fidelity, Aggregation of representations	Loss of sense of individual presence hampers informal interaction.
Printed cardboard cut-outs	Motion of represented participant, Indication of focus-of-attention	Live image of remote user is critical.
Web-based video chat kiosk	Ability to address single person, Personal conversation space	Web-based video chat is optimal for informal 1-1 conversations, but fails when there is a larger group at one end of the connection.
Turn-to-speak indicators	Overlap in turn-taking to speak, Ability to attract others' attention	Group norms affect the adoption and implementation of a technology.
Microphone ball	Overlap in turn-taking to speak, Aggregation of representations	Making the act of turn-taking explicit by having to pass on a microphone ball introduced a norm of formality in the otherwise informal discussion.
Back-channel chat	Ability to communicate with other remote participants	Creating a shared social identity among remote participants facilitates their participation in the larger community.

1 High-end video-conference system

The first prototype seemed to be the most direct approach and leveraged equipment we had on-hand. We initiated a shared conference with each of our remote participants on a high-end video-conference system. But we found that these systems are optimized for a particular type of many-to-many conversation in which only two ends of the connection include all of the participants. That is, one group in one location speaks only with another group in another location. This creates a natural interaction, where each group is presented as a whole. Gestures, glances and relative position within each group are maintained and provide valuable reference cues for the other group. In the alternative, when several individual participants at several locations are connected to a group, their presentation changes: each is bound within a window that shares the screen with the other individuals-within-windows. As a result, the sense of individual presence and persona that is imparted by a distinct representation or inclusion within a group vanishes.

2 Printed cardboard cut-outs

The first attempt to remediate the loss of individual persona afforded by the presentation of individuals-within-windows was to remove the screen image altogether. We created cardboard cut-outs on which we pasted life-size photographic head-and-shoulder images of each remote participant. The cut-outs were attached to the video-conference system's camera so that when the camera moved, the representations moved with it. This provided members of the local group with an 'at-a-glance' indication of what remote participants could see, and through that, where their attention was focused. But the provision of focus-of-attention information provided by the moving cut-outs was not perceived by local participants as compellingly real as the shared video representations.

3 Web-based video chat kiosk

The second attempt to remediate the loss of individual persona was more successful. We initiated one or more chat sessions with remote individual participants using free Web-based video chat applications. Each session was conducted on a separate laptop and each remote participant had a distinct space within the local room. Eventually, we placed the laptops atop kiosks made of foam

cubes that could be adjusted to the height of a standing or seated person. In this way, the remote participant would be at eye-level during both informal socializing and more formal presentation.

In keeping with the way we found that high-end video-conference systems are optimized for many-to-many conversations, we found that Web-based video chat is best suited for close-proximity one-to-one conversations, and laptop onboard cameras and microphones are optimized to be used this way. When one end of the connection is a group, rather than an individual, the interaction suffers. Because remote participants may need to see whiteboards or video screens or local participants who are across a room, a higher resolution camera, like that used on a high-end video-conference system, is needed. And because the primary speaker may also be located across the room, or several local participants may speak in succession, lapel or ambient microphones are also needed. Taken together – a video image (face) in one location, high-resolution camera (eyes) in another, and specialty microphones (ears) spread around the room – remote participants become 'disembodied,' and local participants don't know which direction to look or speak when addressing them.

At times, when a remote participant's audio would be active but his video image was unavailable, we would position his laptop as usual and put his static image on the screen. However, this introduced what we call the 'deception of the static image.' From across the room, it appears that the remote participant is actively participating from his laptop kiosk. Local participants would approach and speak, only to find that the image was static, and the remote participant was not responsive. This happened because of the disembodiment: he was viewing another part of the room through the high-definition camera, not the one located on the laptop, and because there are several microphones, could not tell that it was he (at his laptop kiosk) being addressed.

4 Turn-to-speak indicators

With distinct physical embodiments for remote participants in the local space, it seemed natural to try to provide further capabilities afforded by that physicality. We separately tested two indicators used to indicate a request to speak in turn. One was a mannequin arm and hand attached to the side of a kiosk that could be raised or lowered; the other was an incandescent light bulb positioned immediately above and behind the kiosk. Both were actually operated manually by a facilitator in the local space, who was prompted by an instant message from the remote participant, but the appearance to the rest of the group was that the remote participant had initiated the action. Neither prototype was met with any success, which we believe can be attributed to a few reasons. First, they were not particularly visible, since remote participant kiosks were positioned around the perimeter of the room to keep their power, microphone and camera wires clear of the main walking areas. Second, it is generally not the social norm of the group to raise hands to request to speak up – one just speaks up. Third, being a remote participant creates a limited field of view and hearing. Accordingly, some of the action within the local group is missed. The resulting uncertainty creates a reluctance to speak up over concern for interrupting another speaker that is not seen or heard.

5 Microphone ball

We then introduced a wireless microphone, which was housed within a foam ball and passed from one local participant to another in order to speak (other than the presenter, who wore a lapel microphone). The microphone ball was employed primarily during the Q&A. It represented the ability of the remote participants to hear, and therefore served as a focal point for attention and turn-taking. In that role, it presented them to the local group as more of a community than their otherwise distributed presence.

A number of times, a speaker would forget to 'talk to the ball' and the community would remind him of the oversight. Thus, while there previously was no formal norm for taking turns during Q&A, the microphone ball introduced a protocol that carried with it a greater sensitivity to remote participants in addition to a greater degree of formality. The ball slowed the otherwise informal, rapid pace of discussion in a way that the turn-to-speak indicators could not, and to the benefit of the listeners.

Another unanticipated benefit was that the act of passing the ball from person to person created a soft whoosh sound that was audible to remote participants. This provided them with a signal that a change in speaker was about to occur, a sense-making capability that was previously quite limited, and which helped to draw them into the conversation.

6 Back-channel chat

The feature that met with most appreciation from remote participants grew out of the need for a local mediator to adjust microphone levels, raise and lower kiosk heights and actuate mannequin arms. By joining all the remote participants in a single chat room, the mediator created an environment where they could greet and speak to each other, ask questions, request clarification of a point the speaker made, or discuss other topics. The resulting 'water cooler chat' fostered the creation of a group social identity, separate from that of the local and remote participants as a whole. It was their place to catch up with other similarly sensory impaired participants, who shared their perspective over the flow of sessions.

We encouraged local participants to join the back-channel chat by announcing its activation, posting the instant messaging identities of remote participants on their kiosks, and projecting the real-time chat log onto a large display within the room. Yet few ever participated. One explanation may be that the chat was active primarily during the presentation portion of sessions, a time when local audience conversation was already infrequent, so they may not have felt the need for increased communication. Alternatively, the availability of an immediate neighbor meant that they could speak softly or ask for clarification without the need for an alternative channel to do so. A final explanation may be that a significant proportion of local participants were not well acquainted with the remote participants, and found it difficult to initiate and build congenial relationships without any initial 'face-to-face' introductions [9].

Figure 2 – The most recent version of our problem-finding explorations shows two remote participants, connected via Web-based video chat, speaking with a local participant. Their screen images are positioned atop kiosks at eye-level, and each has a local camera and

speaker. Their names, location in the world, and instant messaging identity are shown so local participants can join them in the back-channel chat. The foam microphone ball, which provides a single focus for all speakers during discussion, is visible at the bottom left.

4 REQUIREMENTS TO IMPROVE COMMUNICATION

Our explorations, and the insights they provided, lead us to develop a set of requirements for the design of 1-to-1-to-many communication setups, with particular emphasis on transitioning between levels of informality in the social context.

1 Provide a shared communication back-channel to foster a social identity.
Remote participants experience impaired sensing capabilities relative to local participants, and they share this characteristic. It creates an identity that is distinct from the main group. A back-channel chat room allows them to talk with each other about the content of the discussion, ask for clarifications, share their misunderstandings and frustrations, or wander off to other topics. Furthermore, the ability to talk without having to whisper or be hushed by a neighbor may promote an even greater freedom of expression, particularly during formal exchanges, than that enjoyed by local participants. In this way, the mechanism for meeting the needs of remote participants provides a capability beyond that of even 'being there' [10].

2 Use microphones that transmit personal, rather than ambient, audio.
Microphones should be arranged so that personal speech, whether by a nearby colleague or a speaker across the room, can be clearly understood. A lapel microphone generally stays with a single speaker, a handheld microphone is passed between speakers, while a boom microphone with sufficiently narrow audio field can be centrally located. Ambient audio works well in contexts where a single person speaks at a time or in conversation at very close range, but it is not well suited to contexts where there are nearby concurrent informal conversations. Listeners hear a 'wall of sound' and cannot distinguish comments intended for anyone in particular. These arrangements require agreement among the local participants to accommodate those who are at a distance. Members of our group pass and speak into the microphone ball even for a single remote listener. At this point, the norm has become established.

3 Video cameras should be able to provide for context-switching.
Each of the three categories of social exchange identified implies a different view of the local space. Congenial conversation occurs within arm's length and requires a face-to-face viewpoint; making sense of a presenter and her displayed material requires the ability to resolve great detail from across a room; and active participation in a group discussion requires the ability to scan the breadth of the gathering of local participants. Video cameras to provide all three capabilities are required, whether they are realized through high-definition cameras collocated with each individual's representation – perhaps the ideal arrangement for preserving the perception of individual presence – or through low-resolution Web cameras at each station with a single high-quality camera positioned for a clear view. The latter arrangement worked for designX because the need for high resolution was limited primarily to the structured presentation and remote participants were sent copies of presentation materials beforehand.

4 Separate representations of remote participants for informal interactions.
During a presentation, the audience – whether local or remote – acts as a collective single. It does not matter much if members have individual or aggregate representations. But it is critical to be presented as a separate individual for person-to-person interaction, which occurs most commonly during social settings or small-group design sessions. Video conference systems aggregate remote participants at distinct locations onto a shared display, so they are less well-suited to maintaining individual persona during informal interactions as multiple instances of Web chat running on separate displays.

5 The technology support activity should not obstruct the mediator's participation.
At all sessions, remote participants are connected via the back-channel to one or two local mediators, whose role is to support the communication technology in their local space. They take

care of kiosk setup and breakdown, replace microphone batteries and arrange for presentation materials to be sent to remote participants. But the role of mediating is sufficiently consuming that at most sessions, mediators missed much of the presentation and dialog content and were not active contributors. This indicates that our current approach requires too much effort for mediation, and that more autonomous designs for communicating with remote participants are required. Our future prototypes will focus on this need.

5 SUMMARY

Addressing the communicative needs of remote participants as a part of design research collaboration while remaining sensitive to the needs of the local group requires a combined technological and social approach. The problem-finding explorations presented in this paper have led us to identify specific requirements for meeting these needs, and for fostering a group social identity, while at the same time preserving individual identities for informal participation. Prototypes such as the microphone ball, back-channel chat and Web-based video kiosks are first steps in designing a technology to satisfy these requirements. Implementing such a communication technology in settings that transition between levels of social exchange would lower the barriers to participation in a distributed community of practice, and encourage collaboration at a more informal and organic level.

REFERENCES

[1] Lave, J. and E. Wenger. *Situated learning: legitimate peripheral participation*, 1991 (Cambridge University Press).

[2] Hildreth, P., Kimble, C. and P. Wright. Communities of practice in the distributed international environment. *Journal of Knowledge Management* Vol. 4, 2000, Pg. 27-37.

[3] Squire, K. and C. Johnson. Supporting distributed communities of practice with interactive television. *Educational Technology Research and Development* Vol. 48, 2000, Pg. 23-43.

[4] Dourish, P. and S. Bly. Portholes: supporting awareness in distributed work group. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Monterey CA, 1992, Pg. 541-547.

[5] Nass, C. and L. Mason. On the study of technology and task: a variable-based approach. in Fulk, J. and Steinfeld, C. eds. Organization and Communication Technology, 1990, Pg. 46-67 (Sage, Newbury Park).

[6] Clark, H. *Using language*. 1996 (Cambridge University Press).

[7] Olson, G. and J. Olson. Distance Matters. *Human Computer Interaction* Vol. 15, 2000, Pg. 139-178.

[8] Lee, K. Presence, explicated. *Communication Theory* Vol. 14/1, 2004, Pg. 27-50.

[9] Rocco, E. Trust breaks down in electronic contexts but can be repaired by some initial face-to-face contact. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Los Angeles CA, 1998, Pg. 496-502.

[10] Hollan, J. and S. Stornetta. Beyond being there. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Monterey CA, 1992, Pg. 119-125.

Contact: David Sirkin
Stanford University
Department of Mechanical Engineering
424 Panama Mall
Stanford, CA 94305
USA
(650) 723-9233
sirkin@stanford.edu

Researchers David M Sirkin, Neeraj Sonalkar and Malte Jung are PhD candidates at the Center for Design Research in the Department of Mechanical Engineering at Stanford University. Professor Larry J Leifer received his PhD from Stanford University and is Director of the Center for Design Research and the Stanford Learning Laboratory.