Grounding Needs: Achieving Common Ground Via Lightweight Chat In Large, Distributed, Ad-Hoc Groups

Jeremy P. Birnholtz
Thomas A. Finholt
School of Information
University of Michigan
1075 Beal Ave.
Ann Arbor, MI 48109
{jbirnhol, finholt}@umich.edu

Daniel B. Horn
U. S. Army Research Institute
2511 Jefferson Davis Hwy
Arlington, VA 22202
danhorn@acm.org

Sung Joo Bae
Sloan School of Management
Massachusetts Institute of Technology
50 Memorial Dr.
Cambridge, MA 02142
sjbae@mit.edu

ABSTRACT
This paper reports on the emergent use of lightweight text chat to provide important grounding and facilitation information in a large, distributed, ad-hoc group of researchers participating in a live experiment. The success of chat in this setting suggests a critical re-examination and extension of Clark and Brennan's work on grounding in communication. Specifically, it is argued that there are some settings characterized by reduced information and clarification needs, where the use of extremely lightweight tools (such as basic text chat) can be sufficient for achieving common ground – even when conversational participants are unknown to each other. Theoretical and design implications are then presented.

ACM Classification
H5.3. Information interfaces and presentation (e.g., HCI): Group and Organization Interfaces

Keywords
Collaboratories, common ground, cyberinfrastructure, cyberscience, chat, instant messaging, distributed groups.

INTRODUCTION AND BACKGROUND
Clark and Brennan [8] define common ground in communication as a state of mutual understanding among conversational participants about what it is that is being discussed. Their widely cited work on the use of various communications media in achieving common ground can be interpreted to mean that popular computer-based collaboration tools, such as chat and instant messaging (IM), should have significant liabilities in terms of supporting collaboration in ad hoc groups (i.e., groups that have no prior history of working together). Specifically, chat and IM lack key attributes, such as co-presence, visibility, and audibility, that Clark and Brennan claim can be necessary for communicators unknown to each other to develop mutual understanding. Recent literature on the use of chat and IM at work and in social settings offers support for Clark and Brennan's claim. For example, the review below suggests that chat and IM appear to be most successful when used within existing groups, such as standing work teams or circles of friends, where the impoverished nature of chat and IM is less of a liability due to familiarity and experience with the other communicators.

Consistent with these findings, we would expect chat and IM to fail when users do not share significant pre-existing common ground or a shared reference frame– such as members of groups that have not worked together before. As the case study below shows, however, the ability of chat and IM to support construction of common ground may be less constrained by features of communication media than Clark and Brennan's theory would predict. In particular, the capacity of a particular communication medium to support construction of common ground may be less a function of characteristics of the medium (e.g., visibility and etc.) – and more a function of communication strategies used within the medium. The paper that follows develops this argument through a review of current research on chat and IM, elaboration of a theoretical perspective that casts opportunities for creation of common ground in terms of communication strategies, and then demonstration of the theoretical perspective via a case study of chat use within a large, geographically distributed ad hoc group.

Chat and instant messaging
Chat and instant messaging (IM) have been studied in both work and social settings. In the workplace, much of this work has been concentrated in the following areas: adoption challenges [20], usage patterns and tendencies [19, 22, 23], with a few studies exploring design issues, such as the value of persistent text in establishing common ground in small groups engaged in tightly coupled collaborative tasks [15].

In the social arena, [17, 18] examine the general character of IM in the life of teenaged subjects, while [16, 29] examine the use of Short Message Service (SMS) and other technologies in social groups.

One underlying theme in these studies is the use of IM and chat to provide "lightweight" support for spontaneous interaction between groups of individuals known to each other. These studies leave open, however, the question of whether chat and IM are similarly effective in settings where participants are less familiar. Examples of such settings include the topic-oriented chat rooms provided by America Online and by Internet Relay Chat (IRC); or the setting described here where remote participants are interested in observing a specific activity.

Awareness

There is a growing literature focused on "fixing" problems with chat and IM. A principal theme of this research is that the impoverished nature of chat and IM (in terms of Clark and Brennan's media traits) must be augmented through various means to make these media more useful for communication (i.e., by reducing the effort required to achieve common ground). Such features are presumed to be particularly useful in the provision of a shared reference framework for common objects in a virtual space, or for understanding, by proximity, who is paying attention to whom or what. For example, one popular approach has been to create proxies that simulate the function of visual awareness in a collocated setting, such as the Babble system [13], Chat Circles [32] or V-Chat, which makes use of elaborate graphical avatars [28]. Such visual awareness provides a very basic simulation of visibility and co-presence in a shared space.

Audio cues have also been used to simulate awareness and audibility. These have ranged from simple "sound IDs" that represent specific individuals in a computer-mediated communication space [22], to various implementations of full-duplex audio, such as "voice loops," that link groups of individuals in real time [21, 33]. The advantage of voice loops is that a very high-bandwidth link provides a great deal of information quickly – and for people who have been continuously participating – a sense of how common ground has been achieved and maintained. However, voice loops aren't nearly as beneficial for people dropping in and out of communication, primarily because there is no persistence. That is, a participant who steps away for a moment or who arrives late must ask for a repetition of previously stated material. As with the general discussion of chat and IM, the open issue with respect to augmenting chat and IM to increase awareness is in understanding what benefits these features provide to users in cases where participants are not well known to each other.

EXTENDING CLARK AND BRENNAN'S PERPSECTIVE

Clark and Brennan [8] present eight properties of media that act as constraints on the grounding process: copresence, visibility, audibility, cotemporality, simultaneity, sequentiality, reviewability, and revisability. In their various combinations, the presence or absence of these constraints can affect the cost of grounding to participants in a communication setting.

One unfortunate side effect of this enumeration of constraints is a frequent misinterpretation suggesting that media with fewer constraints are always preferable in settings where participants have little pre-existing common ground. In some situations, this is true. Experiments where participants from different cultural or language backgrounds must complete a collaborative task, for example, suggest that video and audio channels help in negotiating common ground [31].

We contend, however, that there are many settings where this lesson does not hold true, such as those where participants have substantially different information endowments. For instance, one participant may have more information about a subject of conversation than other participants.

In Clark's [7] experiments, the common ground negotiated by participants involved a shared frame of reference. In the shape experiments, for example, the two participants were presented with separate, but identical, sets of shapes and had to negotiate a commonly understood way to refer to those shapes. In this context, it was apparent that co-presence, visibility, and audibility made a difference in the participants' ability to negotiate understanding.

Imagine a roughly analogous experiment, though, where only *one* participant has a set of shapes that the pair must interact with together. In this case, the player with the shapes can, essentially, dictate the shared reference frame. She might do so by simply providing the other participant with a list of shapes and ways to refer to them ("one looks like a cat, the next looks like a snowflake, etc."). In this case, the most important constraint on grounding is arguably reviewability. The participant without the shapes must be able to refer back to the list provided by the shape-holding participant, and they are subsequently able to discuss the shapes in a meaningfully grounded way. The essential point here is that the difference in the information (in this case about the nature of the shapes being manipulated) possessed by the two participants vastly simplifies the negotiation of common ground, with the implication that a highly constrained medium (i.e., deficient on some or all of Clark and Brennan's properties) could work quite well even with participants unknown to each other. It is important to note that "information" is here being used in an extremely broad sense. It could be the case that one participant has substantially more knowledge or experience, has some sort of privileged access to an event taking place that other participants do not have, or simply has more contextual information (as in the hypothetical experiment above).

This difference in information results in what we might call reduced *grounding needs*. We use *grounding needs* to refer to the amount of ambiguity that must be resolved in the negotiation of common ground. In other words, Clark's original experiment has high grounding needs because both participants were presented with identical sets of ambiguous shapes for which they had to negotiate a shared reference frame. In the hypothetical experiment presented here, however, the amount of ambiguity to be resolved is drastically reduced by the fact that the one participant has no shapes. Thus, the grounding needs are lower because the negotiation of common ground is simply a matter of making a list that both participants can refer to.

Grounding revisited

Clark and Brennan's work refers primarily to settings in which the participants have approximately equal access to information, but we argue that grounding is an equally important process where this is not the case.

Consider carefully the grounding criterion posed by Clark and Brennan: "The contributor and his or her partners mutually believe that the partners have understood what the contributor meant to a criterion sufficient for current purposes" [8] p. 128. Grounding, then, is the "collective process by which participants try to reach this mutual belief" [8] p. 128. The process of grounding occurs via a sequence of presentations and acceptances.

One key aspect of the grounding criterion for our discussion here is the "sufficient for current purposes" phrase. In other words, participants do not necessarily need a complete understanding of what is going on—indeed in the hypothetical shape experiment, the participants without the shapes can arguably never have a complete understanding of the shapes. Rather, they merely need an understanding that is sufficient for the task at hand. In the types of situations we are discussing, this may simply consist of a single presentation ("here is a list of what the shapes look like") and a single acceptance ("sounds good to me").

Toward a theoretical extension

Our essential point is that situational characteristics can influence the (conscious or not) choice of particular communications strategies. It is the constraints of these strategies, and not just the media used, that influence participants' ability to negotiate common ground successfully.

What, then, do we mean by communication strategy, and what characterizes different communication strategies? In considering the answers to these questions, let us first draw on Clark and Brennan's "principle of least collaborative effort." This principle suggests that in a conversation, participants seek to reduce the amount of effort that each must put into formulating new contributions, and interpreting others' contributions. From this, we can derive two defining characteristics of a communication strategy:

1. **How much information must be exchanged?** Some situations require the exchange of very detailed information about the exact nature of events or common points of reference. Such interactions frequently require the exchange of large amounts of information, some of which may be ambiguous and require clarification. This leads to point number two.

2. **How much clarification is necessary?** Where complex, detailed, or ambiguous information is to be shared, there may be need for extensive clarification and discussion on the part of some participants. In these settings, it is also more likely that the contributor will be unsure of whether or not the other participants understood her contribution, and therefore want acknowledgement of their understanding.

From these two characteristics, we derive Table 1, which provides examples of settings where different strategies might be appropriate:

		Information volume	
		Low	**High**
Need for Clarification	**Low**	Distributed participation in a live event (e.g. MOST experiment)	Group task in a shared space with shared objects (e.g. assembling a jigsaw puzzle)
	High	Activities involving potential danger (e.g. launching weapons)	Distributed group task with ambiguous objects (e.g. Clark shape experiments) or possible danger (e.g. telemedicine)

Table 1 Examples of settings characterized by different information volume and clarification requirements

As we will show in the case study presented here, we witnessed the emergence of a low information/low clarification strategy that proved quite successful in grounding a large ad hoc group via text chat.

RESEARCH CONTEXT

An arena likely to produce collaboration where common ground will be difficult to achieve is the emerging area of "cyberscience,"[24] or scientific and engineering research conducted within geographically dispersed, multidisciplinary teams using high performance computers and networks. For example, over the past decade significant interest has focused on the development of collaboratories to support resource and data sharing by scientists and engineers [2, 4, 10, 11, 14]. In particular, a recent report commissioned by the U.S. National Science Foundation [3] recommended an annual investment of US $1 Billion in the creation and development of "cyberinfrastructure," or the collection of services, computational resources, and networks that enable collaboratories and other knowledge environments. One example where cyberinfrastructure is being used to support engineering research is NEESGrid, the collaboratory component of the NSF's George E. Brown, Jr. Network for Earthquake Engineering Simulation (NEES).

The goal of NEES is to broaden participation in earthquake engineering, to accelerate the pace of earthquake engineering research, and to move research findings more quickly into practice to reduce loss of life and property due to earthquakes. NEES will accomplish these goals through an $81.9 million, three year program to improve seismic testing and simulation equipment at fifteen universities (NEES Equipment Sites) and construction of the NEESgrid collaboratory to enable network-based access to equipment and data for researchers at other institutions.

Earthquake engineering (EE) research is concerned with understanding the responses of materials and structures to seismic forces. Work consists of numerical simulation, laboratory tests, and field evaluation of structures and soils. For example, in a structural lab test, a full-size or scale model of a real-world structure is constructed, instrumented with sensors and placed on a large testing apparatus in the laboratory, such as a shake table or reaction wall. This specimen is then subjected to a series of increasing stresses until it experiences structural failure [27]. Due to the scale and cost of earthquake engineering lab equipment, there are a limited number of facilities where earthquake engineering research can be conducted. Therefore, the NEESgrid collaboratory plays an important role because it creates a way for students, faculty, and practitioners without local labs to participate in earthquake engineering research (e.g., through teleobservation of experiments and so forth).

Another possibility that arises from linking experimental facilities via NEESgrid is that these facilities can now be run together to conduct integrated experiments. To understand this concept, imagine a test of a highway bridge. One site may test a physical model of a bridge bent, or supporting archway. Another site may test a physical model of a bridge column. And a third site may run a numerical simulation of the bridge deck. By joining these activities via NEESgrid, feedback and instructions can flow between the three separate simulations (one numerical and two physical) as if they were a single structure. However, since there is no precedent for this type of test in the earthquake engineering community – and because such a test brings together researchers from different institutions and different intellectual traditions (e.g., physical vs. numerical simulation) – there are significant challenges to creating and maintaining common ground. The next section describes the setting and participants in an early integrated test conducted via NEESgrid in July 2003.

METHODOLOGY

Following an iterative, spiral development model [5], NEESgrid has evolved through a series of realistic trial deployments where users employed the NEESgrid software to conduct research tasks. The July 2003 integrated test described here was the most ambitious deployment to date in the NEESgrid project, which started in September 2001 and was scheduled to deliver the final NEESgrid system by October 2004.

The MOST experiment

The Multi-site Online Simulation Test (MOST) took place on July 30, 2003 and involved physical testing apparatus at the University of Illinois at Urbana-Champaign and at the University of Colorado at Boulder, and a numerical simulation run at the National Center for Supercomputing Applications (NCSA). Figure 1 illustrates the experimental setup. A combined team of earthquake engineers and NEESgrid system developers devised the experiment. It involved the simulation of a two-bay single-story steel frame structure, with one part physically in Boulder, a second part physically in Urbana-Champaign – and the connection between the two simulated by the computer model at NCSA. This experimental setup provided an opportunity to showcase, test and critique critical aspects of the NEESgrid collaboratory, including: grid capabilities, data management, telepresence (including remote control of experimental equipment and the ability for remote participants to chat with each other and observe live video and data feeds from the experiment), data acquisition and streaming services, and computation and simulation. This paper reports primarily on user experience with the chat component.

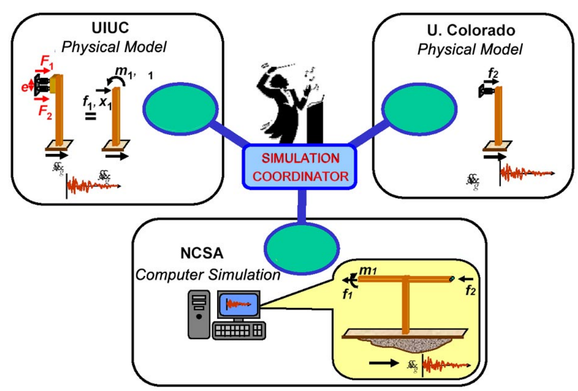

Figure 1 A diagram of the three components of the NEES experiment at UIUC, Colorado and NCSA.

There was intense interest in the MOST experiment throughout the EE research community. It is important to note, however, that most of the individuals involved were largely unknown to each other and for most this was to be their first encounter with the NEESgrid interface.

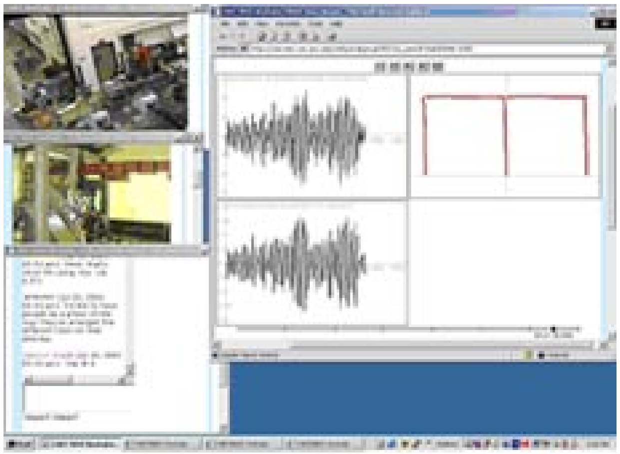

Figure 2 One user's configuration of the NEESgrid interface. Note video feeds at the top left, streaming data views on the right, and chat in the lower left.

The experiment ran from approximately 11:00 AM (all times are CDT) until 6:30 PM. During this time users could log in to the web-based NEESgrid interface, and access the telepresence components of the experiment. Specifically, this included the ability to view data from the experiment, feeds from video cameras at the two equipment sites, and to chat with other participants. Using NEESgrid, users were able to individually select and configure a range of views that included any or all of the cameras and sensors in use during the experiment. Figure 2 shows a screen shot of one user's configuration of this interface during the experiment. At the peak of participation, there were a total of 65 unique users logged into the server simultaneously. A total of 39 participants logged in from NEES equipment sites. Server records of logins, chat usage and chat content were kept and analyzed.

In addition, participants were asked to complete two short online surveys before and after the experiment. The pre-experiment questionnaire was administered one week prior to the experiment to members of the equipment site and system developer communities and asked about their expectations for the experiment. The post-experiment questionnaire was administered during the week following the experiment, and asked about the actual experience relative to expectations. Both instruments asked for free-response comments. 77 individuals completed the pre-experiment questionnaire, and 22 individuals completed the post-experiment questionnaire.

RESULTS: WHAT HAPPENED?

This section of the paper describes the outcomes of the MOST experiment, looking first at user experience and then focusing more clearly on the communication strategies that emerged over the course of the chat interaction.

User experience

It is important for several reasons to note that the experiment was widely considered to be a success. In the first place, this validates the development effort and provides feedback for future development on the NEESgrid components listed above. Second, and more important for our purposes here, it strongly suggests that remote users did have a sense of what was going on and that the grounding we discuss in the next section was actually achieved.

Participants were asked before the experiment about their expectations for the performance of the NEESgrid system. Specifically, they were asked whether they thought the experiment would have "no defects," "a few minor defects, and no major defects," "several minor defects, a few major defects," or "several minor and major defects." On the whole, response to this item can be characterized as cautiously optimistic, with 5% expecting no defects, 68% expecting a few minor defects, 25% expecting several minor and a few major defects, and 1% expecting several minor and major defects (n=77).

Actual performance was assessed after the conclusion of the experiment, and response was generally quite positive. On a 5-point Likert scale anchored by "very successful" and "very unsuccessful", 64% of respondents ranked the experiment "very successful" or "successful," 27% were neutral, and 9% ranked it "unsuccessful" or "very unsuccessful" (n=22). In addition, in a "quick response" survey immediately following the experiment, 60% of respondents indicated agreement or strong agreement with the statement that "Overall, the NEESgrid system was excellent during the MOST experiment on July 30, 2003."

Chat Usage

As we have noted previously, the chat feature emerged as critical in providing a grounding framework for all experiment participants. This part of the paper examines the usage of chat in three ways: 1) usage statistics, 2) analysis of content, and 3) a closer look at the role of chat in grounding.

Usage statistics

Let us first consider the volume of usage of the chat feature during this experiment. There were a total of 532 chat utterances during the course of the experiment. The rate at which messages were sent was somewhat "bursty," with the highest volume of messages being sent earlier in the day, in the middle of the experiment, and at the end of the day. As we will discuss later, this makes sense in light of the grounding role that the chat feature played.

Of the 152 who participated in the experiment, 35 individuals contributed to the chat stream at least once. We were interested in who these users were and what they had to say. As we mentioned earlier, both NEESgrid system developers and earthquake engineers were participating in the MOST experiment. We therefore coded the contributions to the chat stream based on the program role of the contributors. As Figure 3 shows, the majority of the messages (340) came from the System Integration (SI) team members, while 171 came from users at the NEES equipment sites (ES). This is quite interesting in that most of the participants in the experiment were earthquake engineers.

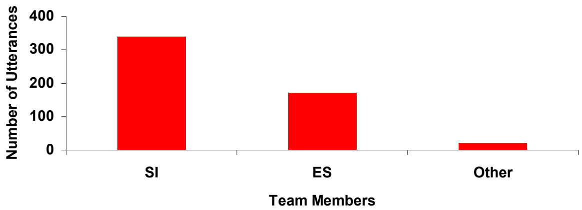

Figure 3 Number of chat utterances that came from System Integration (SI) team members, Equipment Site (ES) team members, and others.

In addition, we found that of the 340 messages from the SI group, 229 of them came from a single user. In other words, this user was responsible for 43% of the total chat stream. The next sub-section explores the nature of the chat content, with a focus on this participant's dominance – and what it revealed about the emergent communication strategy the MOST experiment group adopted.

Chat content

In an effort to better understand the usage of chat during the MOST experiment, we read and coded each of the 532 utterances into 13 relevant categories. Two individuals coded the messages, and Cohen's kappa was calculated to be .6, which indicates a "good" level of inter-coder agreement [9]. Of these categories, the four most common accounted for 79% of the messages and included:

1. **Grounding/orienting statements (45% of total)**: These were generic statements (i.e. not addressed to a specific user or question) that served to inform participants what was going on at the physical experiment sites, where they could find needed information, how they could access different components of the experiment, etc.

2. **Troubleshooting questions (19% of total):** These were specific questions from users about different aspects of the interface or specific problems they were having with the system.

3. **Troubleshooting responses (22% of total):** These were responses, usually by the SI team but occasionally by other EE users, to troubleshooting questions. A troubleshooting response usually included a fix for the problem in question, but occasionally asked for more detailed information (e.g. "What browser are you using?").

4. **Continuation of conversation (10% of total):** In cases where a troubleshooting conversation went beyond a single question and response (such as when the questioner said "That worked, thanks!" subsequent utterances were coded as "continuing conversations." This code was also used for continuing conversations in the categories described below that are not elaborated here.

There are several interesting things to note here. First is that, as we might expect, about 63% (categories 2-4, plus remaining categories) of the chat messages were concerned with the usual content of chat—resolving specific questions, coordinating small groups, and general social pleasantries. This is what chat tools are generally designed to support, so it is not surprising.

More interesting and important for our purposes, though, is that 195 of the 532 total messages (37%) contained content related primarily to grounding and orientation. This was not what the tool was designed for, but it emerged as a critical aspect of the experiment. Additionally, members of the system development team sent all of these grounding/orienting messages and 144 of them (74%) came from the single user mentioned previously. Thus, it seems clear that this user was playing a dominant role in helping participants understand what was going on in the experiment. What he or she was saying and how it helped then become important questions, and these are addressed in the next sub-section.

Grounding and orientation via chat

We divided the 195 grounding messages into four sub-categories, and then described the on-the-fly development of communication protocols that went with each sub-category. The sub-categories and protocols are described here and examples are provided:

1. **Physical site updates:** Site updates summarized the status of activity at the two equipment sites (Illinois and Colorado). The updates were intended to let remote participants know what was going on, particularly during delays. In the early part of the day, these appeared as an undifferentiated part of the chat stream, such as this early status update:

   ```
   (Jul 30, 2003 11:06 am): Today's experiment is not yet running. We hope to be able to start within 30 minutes. I will provide updates as needed
   ```

   Later, however, an informal protocol was adopted whereby updates of this nature were preceded by the word "UPDATE." The first instance of this was at 11:32 am, though the example below is a bit later:

   ```
   (Jul 30, 2003 12:25 pm): UPDATE: If you go to the MOST "Data Viewer" section now, you will see a selection for "MOST experiment underway (stored data)". This is the data coming from the current experiment run (see above).
   ```

2. **Experiment timing updates:** Earthquake engineering experiments are run in a series of finite time-steps, and these time steps are used to sequence and synchronize data streams from various sensors and cameras.

Because the data being released on the web were delayed slightly, it was necessary to broadcast the current timestep to all remote participants, as seen here. At first, several remote users asked for this information and these requests were heeded. Eventually, one of the system developers took to sharing the current timestep on a regular basis, as seen here:

> (Jul 30, 2003 02:04 pm): We are at timestep 400 now.

3. **Instructions for newcomers:** As the experiment took place over several hours, remote participants came and went throughout the day. Thus, it was important to provide brief, but coherent instructions to newcomers and to returnees. At first, users asked many questions about how to accomplish specific tasks. As the day went on, we saw three occasions where one SI team member spontaneously broadcast a whole series of instructions for newcomers:

> (Jul 30, 2003 02:19 pm): I'm going to run through the "newbie" instructions again for those who've joined recently.
>
> (Jul 30, 2003 02:19 pm): We are currently somewhere around timestep 500 in a 1500 timestep experiment run.
>
> (Jul 30, 2003 02:20 pm): You can see the physical experiment sites via telepresence cameras (UIUC and Colorado)…

4. **Technical orientation:** Finally, there were several occasions where trouble with the NEESgrid system itself had to be explained. In one case, for example, the server ran out of disk space and was rebooted:

> (Jul 30, 2003 01:25 pm): The experiment run has continued, despite the web interface restart, so we should shortly be seeing the data continuing in process...

Through these orientation messages, remote participants were provided with a wealth of information throughout the duration of the experiment. One remaining question, however, is whether or not this information was useful and adequate. Our results suggest that it was both.

First, Figure 4 shows that the peak in orientation/grounding messages occurred around 2:00 pm. After this time, the number of messages in the other categories began to drop, with troubleshooting questions dropping particularly steeply. This suggests that people felt better oriented later in the experiment and that the orienting messages may have helped in this process – although learning effects can't be ruled out as an alternative explanation.

Supporting the claim that the orienting messages had impact were comments from several users who cited the chat tool when asked about the most useful aspect of the NEESgrid interface. Others, who never found the chat feature in the interface reported feeling particularly disoriented and out of touch.

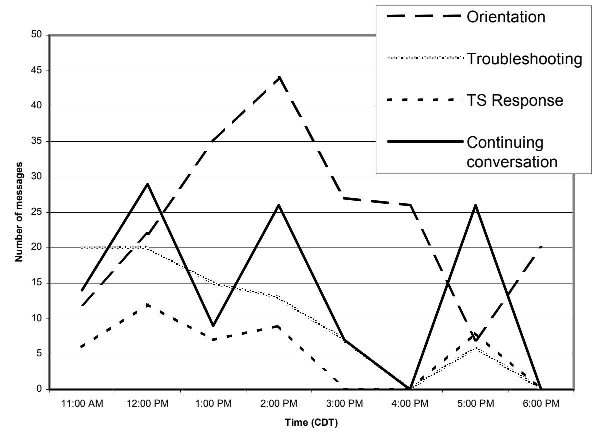

Figure 4 Number of chat messages of each type per hour prior to and during the experiment

DISCUSSION

The usage of chat as a grounding and orientation tool raises several important issues for discussion that have largely been absent from the literature on chat and IM technologies.

Theoretical implications

In the MOST experiment we observed the emergence of a particular communication strategy in a setting characterized by low information and low clarification requirements.

In terms of information, remote participants were like the participants without shapes in the hypothetical experiment described earlier. They had very little sense of what was taking place at the remote sites, and needed only a basic sense of this in order to feel adequately grounded as participants in the experiment. Thus, the facilitation information that was shared was relatively unambiguous, basic information about the status of the experiment – and about which tools to use to view the experiment.

In terms of clarification, the relatively low ambiguity of the information being shared combined with the large number of participants rendered clarification generally unnecessary. Moreover, the practice that appears to have emerged is that orienting utterances were taken to be universally understood unless a participant indicated otherwise. In other words, there was no need for participants to indicate they understood orientating messages (as in ordinary conversation, either explicitly or via nonverbal means).

The combination of low information and clarification requirements led to a situation where text chat, despite the many constraints on grounding that it poses, was an adequate grounding tool. This is similar in important ways to predictions made by adaptive structuration theory [12]. That is, the MOST experiment appeared to be a case of users adapting a technology to suit their needs in a specific situation and achieving highly desirable outcomes.

It might be argued that the low information and clarification needs observed resulted from pre-existing common ground between participants, many of whom had background training in earthquake engineering. While this shared background may have helped some, it was nonetheless the case that this was a novel experimental style that had never been witnessed by any participant, and that a significant fraction of the participants had not met face-to-face prior to this experiment. Thus, some collective negotiation of a shared understanding in a novel setting was required here.

All of this suggests that we must extend our interpretation of Clark and Brennan's work [7] to include not just the attributes of media that facilitate or constrain the grounding process, but also situational traits and communication strategies that may simplify or complicate grounding.

Design and research implications

In the first place, we have mentioned repeatedly that chat and IM technologies have traditionally been designed to support interaction between groups of approximate equals. In the MOST experiment, however, there was a clear disparity between participants who were physically present at one of the experiment sites and those who were participating remotely. Those who were physically present were "privileged" in the sense that they had immediate access to information and expertise about the experiment that purely remote participants did not have. We saw that these participants contributed substantially more to the chat conversation, and that they did so more authoritatively, even developing some informal norms for distinguishing this information from the rest of the chat stream (e.g., using the word "UPDATE"). Moreover, we saw that the remote participants found this information stream valuable in understanding what was going on in the experiment—to the extent that these more authoritative messages served as a sort of "voice of mission control" in providing a common understanding of what was taking place.

This raises several important design and research questions about the use of chat systems for groups of unequal participants. Specifically, these include: 1) means for signaling authority figures, 2), distinguishing between genres of utterances, 3) automating the broadcast of certain technical information, and 4) the importance of facilitation.

Signaling authority

One problem with most chat and instant messaging systems is that it is difficult to distinguish known authority figures from those without authority or even those who might have malicious intent. For example, it would have been quite easy in the MOST experiment for an outsider to log into the system and provide false information. Though some experienced users might know by the name of the contributor that something was awry, a new user would have difficulty making this distinction.

Others experienced a similar problem in developing the chat interface for another collaboratory [25]. In this collaboratory, operators of a data-gathering system in Greenland were provided facilities to chat with researchers using the data at universities around the world, in addition to the general public. The operators reported difficulty in distinguishing comments and suggestions by accredited (legitimate?) researchers from those by "lurkers" from the general public. They requested a means for doing so, such that the input from the researchers could be given higher priority. One interesting aspect of this example is that we see three "classes" of participants—operators with privileged information about the data-gathering instrument, researchers with privileged status in the use and gathering of this data, and the interested, but not privileged, public.

The clear design implication here is that some means is frequently necessary for signaling legitimate privileged status in a diverse chat environment. Some simple ways to accomplish this include color (i.e. displaying privileged utterances in a different color so they stand out) or other visual indicators of status (i.e. displaying privileged text in a different or larger font, etc.). IRC systems, for example, accomplish this by designating operators ("ops") using a special character in the user's name. One could also imagine the placement of privileged information in separate chat streams. For example, in a MOST-like scenario, there could be a separate "Experiment Status" window that includes information like the current time step, and privileged users' comments on the status of the experiment.

Signaling genre

We witnessed the emergence of several genres of chat messages that served orientation and grounding purposes. Again, informal norms were developed on the fly to provide some indication of what was contained in a given utterance or why certain information (e.g. instructions for newbies) was being provided. It would be useful to bear this in mind in designing future systems by providing explicit means for distinguishing between different genres.

Again, color and other visual indicators would be one useful way to solve this problem within the confines of a single chat window. The "experiment status" window suggested above could also accomplish this with certain genres, but one must be careful not to clutter the interface with too many windows, or too many things for a new user to find in order to understand what is going on.

One potential way around this issue would be a user-configurable chat system that initially shows all chat messages in a single stream, but allows advanced users to filter certain genres (or messages from certain people) into separate spaces. The "ignore" or "block" feature on many IM systems does this in a coarse way, but this example suggests that a finer level of control may be more useful.

Automating certain information

Astute readers may have already noted that some of the information broadcast by the system developers in the chat stream need not have come from a human. Many of these messages, for example, announced such things as the availability of an electronic resource, or the current time step of the experiment. This information existed in electronic form, and could potentially have been automatically broadcast to the chat room. For example, as the experiment operators made their data streams publicly available on the web, the server could have triggered an automated chat message indicating their availability.

This is similar to the "email notification" feature provided by many web-based communication and collaboration tools (e.g., WebBoard). Because people typically check their email more regularly than these web sites, it can be useful to receive an email notice when something on the web site changes. Similarly, during the MOST experiment people were watching the chat stream more closely than they were the various other areas of the NEESgrid interface (e.g. the data repository). Thus, "chat notification" emerged on the fly and seems a useful innovation for future systems that operate on a similarly short time horizon.

Additionally, this notification function is similar in some ways to the functionality provided by "bots" in IRC systems. With "bots," users can write programs that provide a great deal of functionality via the chat interface (e.g. checking to see when a user was last seen in a room).

The importance of facilitation

One important lesson from the MOST experience is that facilitation was not only a usage of chat that emerged unexpectedly, but it was also absolutely vital to the success of the experiment. Thus, one key implication from this lesson is that those engaged in similar work should not only design for facilitation and grounding, but also recognize the importance of designating a facilitator.

Additionally, one consistent theme in the negative survey responses was that they tended to be from people who complained about difficulties in finding or using the chat interface. These users not only did not find the experiment successful, but also complained about feeling alienated and disoriented. Thus, people who did not have access to the information provided by the facilitator felt less a part of the experiment—or were less grounded—than those who did.

Broader applicability

Beyond the case of cyberinfrastructure development presented here, there are many other situations where the MOST experience can provide valuable insights. For example, we believe the MOST experiment is a specific instance of a general class of research and education activities that allow for public or "quasi-public" participation via Internet-based technologies.

In science, for example, we see increasing interest in outreach activities that actively engage children in classroom settings, such as Bugscope [26] and Chickscope [6]. There are also web-based video and data streams that allow observation of oceanography research at sea [1].

In another example, telemedicine technologies are currently being used to allow for the training of doctors in Vietnam [30]. Here, remote doctors can view medical procedures, such as open-heart surgery and ultrasound examinations.

What this growing body of examples share is the need to engage remote participants, who may or may not be domain experts, in an ongoing, complicated stream of activity. We contend that the lessons learned in our NEESgrid experience are valuable in these settings.

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation's George E. Brown, Jr. Network for Earthquake Engineering Simulation (CMS #0117853). We wish to thank Mark Handel and Matthew Bietz for their comments on this work, Eric Cook for research assistance, and all members of the NEES team and earthquake engineering community who participated in the MOST experiment.

REFERENCES

1. Explore the Sea. www.explorethesea.com.
2. Arzberger, P. and Finholt, T. *Data and Collaboratories in the Biomedical Community*. National Institues of Health, Ballston, VA, 2002.
3. Atkins, D. E., Droegemeier, K. K., Feldman, S. I., Garcia-Molina, H., Klein, M. L. and Messina, P. *Revolutionizing Science and Engineering through Cyberinfrastructure: Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure*. National Science Foundation, Washington, DC, 2003.
4. Birnholtz, J. P. and Bietz, M. J. Data at Work: Supporting Sharing in Science and Engineering in *Proceedings of ACM Conference on Supporting Group work (GROUP)* (Sanibel Island, FL, November 9 - 12 2003)
5. Boehm, B. W. A Spiral Model of Software Development and Enhancement. in R. M. Baecker, J. Grudin, W. A. S. Buxton and S. Greenberg (eds.) *Human Computer Interaction: Toward the Year 2000*. Morgan Kaufman, San Francisco, CA. 1995
6. Bruce, B. C., Carragher, B., Damon, B. M., Dawson, M. J., Eurell, J. A., Gregory, C. D., Lauterbur, P. C., Marjanovic, M. M., Mason-Fossum, B., Morris, H. D., Potter, C. S. and Thakkar, U. Chickscope: An Interactive Mri Classroom Curriculum Innovation for K-12. *Computers Education 29, 2/3* (1997), 73-87.
7. Clark, H. H. *Arenas of Language Use*. University of Chicago Press, Chicago, IL, 1992.
8. Clark, H. H. and Brennan, S. E. Grounding in Communication. in L. B. Resnick, R. M. Levine and S. D. Teasley (eds.) *Perspectives on Socially Shared Cognition*. American Psychological Association, Washington, DC. 1991

9. Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement 20, 1* (1960), 37-46.

10. Cummings, J. and Kiesler, S. KDI Initiative: Multidisciplinary Scientific Collaborations. http://netvis.mit.edu/papers/NSF_KDI_report.pdf.

11. David, P. A. and Spence, M. *Towards Institutional Infrastructures for E-Science: The Scope of the Challenge*. The Oxford Internet Institute, Oxford, UK, 2003.

12. DeSanctis, G. and Poole, M. S. Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory. *Organization Science 5, 2* (1994), 121-147.

13. Erickson, T., Smith, D., Kellogg, W., Laff, M., Richards, J. and Bradner, E. Socially Translucent Systems: Social Proxies, Persistenct Conversation and the Design of "Babble" in *Proceedings of Computer Human Interaction* (Pittsburgh, PA, 1998) ACM Press, 72-79.

14. Finholt, T. A. Collaboratories as a New Form of Scientific Organization. *Economics of Innovation and New Technologies 12, 1* (2003), 5-25.

15. Gergle, D., Millen, D. R., Kraut, R. and Fussell, S. R. Persistence Matters: Making the Most of Chat in Tightly-Coupled Work in *Proceedings of ACM Conference on Human Factors in Computing Systems* (Vienna Austria, April 24-29 2004) 431-438.

16. Grinter, R. E. and Eldridge, M. A. Y Do Tngrs Luv 2 Txt Msg? in *Proceedings of Seventh European Conference on Computer Supported Cooperative Work.* (Bonn, Germany, September 18-20 2001)

17. Grinter, R. E. and Palen, L. Instant Messaging in Teen Life in *Proceedings of ACM Conference on Computer Supported Cooperative Work* (New Orleans, LA, November 16-20 2002) 21-30.

18. Handel, M. *Communication Technologies and the Freshman Transition: Staying Close with Friends*. Unpublished Dissertation, School of Information, University of Michigan, Ann Arbor, MI, 2004.

19. Handel, M. and Herbsleb, J. What Is Chat Doing in the Workplace? in *Proceedings of Computer Supported Cooperative Work* (New Orleans, LA, 2002) ACM Press, 1-10.

20. Herbsleb, J., Atkins, D. L., Boyer, D. G., Handel, M. and Finholt, T. A. Introducing Instant Messaging and Chat into the Workplace in *Proceedings of Proceedings of ACM Conference on Computer Human Interaction* (Minneapolis, MN, 2002) ACM Press,

21. Hindus, D., Ackerman, M. S., Mainwaring, S. and Starr, B. Thunderwire: A Field Study of an Audio-Only Media Space in *Proceedings of ACM Conference on Computer Supported Cooperative Work* (Cambridge, MA, November 16-20 1996) 238-247.

22. Isaacs, E., Walendowski, A. and Ranganathan, D. Hubbub: A Sound-Enhanced Mobile Instant Messenger That Supports Awareness and Opportunistic Interactions in *Proceedings of ACM Conference on Human Factors in Computing Systems* (Minneapolis, MN, 2002)

23. Nardi, B., Whittaker, S. and Bradner, E. Interaction and Outeraction: Instant Messaging in Action in *Proceedings of ACM Conference on Computer Supported Cooperative Work* (Philadelphia, PA, 2000)

24. Netwich, M. *Cyberscience: Research in the Age of the Internet*. Austrian Academy of Sciences, Vienna, 2003.

25. Olson, G. M., Daniel E. Atkins, Robert Clauer, Thomas A. Finholt, Farnam Jahanian, Timothy L. Killeen, Prakash, A. and Weymouth, T. The Upper Atmospheric Research Collaboratory. *Interactions 5, 3* (1998), 48-55.

26. Potter, C. S., Carragher, B., Caroll, L., Conway, C., Grosser, B., Hanlon, J., Kisseberth, N., Robinson, S., Thakkar, U. and Weber, D. Bugscope: A Practical Approach to Providing Remote Microscopy for Science Education Outreach. *Microscopy and Microanalysis 7,* (2001), 249-252.

27. Sims, B. Concrete Practices: Testing in an Earthquake Engineering Laboratory. *Social Studies of Science 29, 4* (1999), 483-518.

28. Smith, M. A., Farnham, S. D. and Drucker, S. M. The Social Life of Small Graphical Chat Spaces in *Proceedings of ACM Conference on Human Factors in Computing Systems* (The Hague, Netherlands, April 1-6 2000) 462-469.

29. Taylor, A. S. and Harper, R. Age-Old Practices in the 'New World': A Study of Gift-Giving between Teenage Mobile Phone Users in *Proceedings of ACM Conference on Human Factors in Computing Systems* (Minneapolis, MN, April 20-25 2002) 439-446.

30. Uplift International Web Site. www.upliftinternational.org/telemed.htm.

31. Veinott, E., Olson, J. S., Olson, G. M. and Fu, X. Video Helps Remote Work: Speakers Who Need to Negotiate Common Ground Benefit from Seeing Each Other in *Proceedings of ACM conference on Computer-Human Interaction* (Pittsburgh, PA, 1999)

32. Viegas, F. B. and Donath, J. S. Chat Circles in *Proceedings of ACM Conference on Human Factors in Computing Systems* (Pittsburgh, PA, May 15-20 1999) 9-16.

33. Watts, J. C., Woods, D. D., Corban, J. M. and Patterson, E. S. Voice Loops as Cooperative Aids in Space Shuttle Mission Control in *Proceedings of ACM Conference on Computer Supported Cooperative Work* (Cambridge, MA, November 16-20 1996) 48-56.