K. Kuutti, E. H. Karsten, G. Fitzpatrick, P. Dourish, and K. Schmidt (eds.). *Proceedings of the Eighth European Conference on Computer-Supported Cooperative Work, 14-18 September 2003, Helsinki, Finland*

Group-to-Group Distance Collaboration: Examining the "Space Between"

Gloria Mark, Steve Abrams, Nayla Nassif
University of California Irvine, Irvine, California, USA
{gmark,sabrams,nayla}@ics.uci.edu

Organizations are moving towards a new type of work: group-to-group collaboration across distance, supported by technologies that connect rooms across distance into large collaboration spaces. In this study we report on distributed group-to-group collaboration in the domain of space mission design. We use the metaphor of the "space between" distant groups to describe the connections, interdependencies, and gaps that exist. To the extent that the "space between" remains wide, the risk for design errors increases. We found that different teams, who had different processes and methodologies, were able to form hybrid solutions. However, their hybrid solutions addressed mostly terms and results, and did not address the deeper methodologies that created the results. We also found that some individuals acted as information bridges across sites, representing the teams in articulation. To a large extent small groups were used for reconciling perspectives, but the majority of results were not communicated and integrated back into the larger team. We discuss the challenges that group-to-group collaboration designers face in meeting requirements for supporting these new technologies.

Introduction

Over the last decade in CSCW, a number of empirical studies have described the difficulties people face when collaborating at a distance. These studies have mostly highlighted the constraints that exist when people communicate over media channels that limit social information. These studies have typically examined individuals interacting in teams and they span a range of technologies such as an audio media space (Ackerman et al., 1997), desktop conferencing (Mark et al., 1999), a virtual work environment (Fitzpatrick et al., 1996), video (Ruhleder and Jordan, 1999), chat (Bradner et al., 1999), and instant messaging (Nardi et al., 2000) across many domains (Olson and Olson, 2000).

However, in response to demands of combining whole team expertise, organizations are now moving toward a new collaborative configuration: group-to-group collaboration. As opposed to an individual at each site, entire teams are now collaborating across distance in real time. This enables organizations to benefit from larger bodies of increased specialization at distant sites.

To support such large group-to-group collaboration, new technologies are being developed using larger interfaces and multiple wall-sized displays to convey larger views of people and data. Access Grid technologies (2002) and HDTV video conferencing (Mark and Deflorio, 2001) are two examples. Access Grid technologies use interactive multi-media technologies with multicasting, showing multiple views from different sites. There are currently large numbers of site nodes, and these are increasing. The high resolution of HDTV video shown on a wall-sized display is designed to overcome the tradeoff experienced by ISDN video of showing clear talking heads versus the entire room. Despite these exciting technology advances, the study of problems and experiences with group-to-group collaboration still needs more attention. The study of collaboratories (e.g. Finholt, 2003) has been an important step in this direction.

In this paper, we report on distributed group-to-group collaboration in the domain of space mission design. The nature of the task—conceptual space mission design—is highly complex. Design has been characterized as an ill-defined problem (e.g. Carroll 2000). As a result, it requires much negotiation and articulation as design tradeoffs are discussed. Collocated environments are advantageous for designers not only because they provide awareness of the state of the design (Robertson, 1997) but they also enable designers to have immediate access to others, e.g. to question the relevancy of a requirement, to negotiate design tradeoffs, to collectively "walk through" the design and discuss discrepancies. A distributed environment makes these activities more difficult. Furthermore, in an arena with multiple teams and sites, the difficulty of accessing another designer is even more compounded.

Group-to-group and individual distance collaboration are significantly distinct in several ways. First, in group-to-group collaboration there are multiple actors at most nodes, generally teams, whereas in individual distance collaboration, there are mostly individuals at each node, though some nodes may have small groups. Second, group-to-group interaction can be characterized as information being communicated through many different networks, primarily within and across sites. The entire group at each site may participate in information exchange, but also subgroups may interact in parallel. Last, group-to-group technologies to support such collaboration are generally room-sized environments with large data and video displays. These distinctions are not perfect, as for example, in some individual distributed interactions, communication also occurs through multiple networks in parallel.

We envision a new class of interaction problems that exist with group-to-group collaboration compared to distance collaboration with individuals. First, as articulation is a central activity in design, we expect inter-site articulation to be difficult as it involves articulation of entire team perspectives. To add to the difficulty, each team needs to engage in articulation at their own site. Next, we expect that it will be difficult for individual team members to contribute to the discussion simply because of the large number of actors and distance between them. Last, we also expect that information from different actors will be difficult to integrate into the larger team. For example, design tradeoffs may be discussed at one site or across distance by a few individuals, and the results may not get communicated and integrated back into the larger team's design.

The "space between" in group-to-group collaboration

Weick and Sutcliffe (2001) write that high reliability organizations attempt to accomplish on a large scale what people do well on a small scale. A distributed group-to-group space mission design team attempts the same: performing the activities that smaller groups do at one site, but on a larger scale. The coordination increases to manage collaboration not only within, but also between sites. We adopt a relational orientation (Bradbury and Lichtenstein, 2000) which focuses on the relations between individuals and groups in the organization as opposed to focusing solely on the properties of individuals or entities. With this approach, we view a collaborating network as a configuration of relationships and see the space of interaction between team members as where the actual teamwork occurs. A relational orientation enables us to shine the spotlight on the "space between" (Bradbury and Lichtenstein, 2001; Roberts and Yu, 2003), i.e. the connections, interdependencies, and gaps that exist between groups across organizational sites. It is through these gaps where common meaning is lost or misconstrued, and conversely the connections where the potential exists for constructing and/or restoring meaning. We feel that a focus on the "space between" distributed groups enables us to understand where groups are aligned in their perspectives and where they are not. The extent to which coordination and articulation can bridge the gaps determines how much meaning can be transported and aligned across group borders.

The relations among actors have often been focused on in the field of CSCW, using perspectives such as articulation work (e.g. Gerson and Star, 1986), activity theory (Nardi, 1996), and distributed cognition (Hutchins, 1995). We believe our focus on the "space between" distributed groups can inform us of where gaps and connections lie in group-to-group collaboration.

We maintain that, to the extent that the "space between" remains large in collaborating groups, then risk increases for errors in the group product. An example of human and organizational risk in space mission design occurred in the space shuttle Challenger disaster. Vaughan (1997) writes about the "normalization of deviance" in the NASA organization as an explanation for events that led to the disaster. Vaughan's argument points out that human and organizational factors interacted with technology problems to expand the boundaries of acceptable risk. Vaughan showed that the errors occurred in what we call the "space between" distributed groups with different organizational cultures.

We expect the "space between" distributed groups to impact design in different ways. First, articulation has been discussed for some time in CSCW and other fields, e.g. with respect to how people of different social worlds use different terminologies (e.g. Gerson and Star, 1986; Schmidt and Bannon, 1992). However, in group-to-group collaboration, articulation is a much more difficult task than within a single group, as it must reconcile perspectives from entire teams. Teams from various organizational sites can use different terms and even the same terms can be interpreted distinctly in different organizational contexts (Bechky, forthcoming). We expect that team processes are well-known and visible within a team but much less visible outside of a team. Distance impedes the visibility. Second, though current group-to-group collaboration technologies enable a many-to-many information network, we expect that some individuals will act as "bridges" connecting the team sites. In social network theory, information bridges are individuals who have a significant degree of influence on interaction within a network and how information diffuses throughout the network (Freeman, 1991; Bavelas, 1950). The process by which information is conveyed in the large group-to-group configuration can impact the expertise that is communicated from a site. Last, in large groups with semi-formal or informal interaction, it is natural for the team to break up into smaller "sidebar" conversations (i.e. smaller group conversations that are conducted by telephone across sites or by clustering together within a site). Negotiations and decisions reached in some sidebars are important and should be communicated to the larger group. We would expect that multiple sites have difficulties in accessing and integrating the results of sidebar decisions. Though we can focus on many aspects of group-to-group collaboration, as a first step, we examine interaction: articulation, information networks, and information integration.

Research Setting and Methodology

The study took place at a large distributed aerospace organization whose main mission is to research, invent, and develop new technologies to enable (mainly) space-based scientific research. This study examines primarily four engineering teams distributed around the U.S. who synchronously designed an actual conceptual space mission design. Team 1 (Site 1) on the west coast had 24 contributors, Team 2 (Site 2) in the midwest had 12, Team 3 (Site 3) in the south had 9, and a single participant was in the southwest (Site 4). The majority of the people at Site 1 had worked together as a team for several years, and most people at Sites 2 and 3 had years of experience working together within their teams. Site 1 had two facilitators, and Sites 2 and 3 each had one. The purpose of collaboration was to combine different team specializations. Site 1 was responsible for science and mission coordination, and Sites 2 and 3 were responsible for propulsion and power. Sites 3 and 4 shared responsibility for the power supply of the overall system. We refer to all teams together as the entire Design Team.

A combination of technologies was used to support the large-scale interaction. For all distributed design sessions, a shared application, Microsoft NetMeeting, was used on a public display to show linked spreadsheets and graphics that were imported into a software presentation tool. For our primary data set, all four sites were linked by a video-teleconferencing (VTC) service that automatically switched views available to each local site for a short duration. It displayed the video stream from the non-local site that had the greatest average volume. This maximized bandwidth utilization by displaying to one site the most vocally active of the other sites. People were thus subjected to changing views of remote sites that occasionally did not match the locale of the current speaker (i.e. due to delays in volume sampling). Site 1 had three large public displays each 12 x 6 feet (one showed the video and the other two showed NetMeeting), Site 2 had two public displays of 6 x 5 feet showing video and NetMeeting, and Site 3 had one public display of 6 x 5 feet showing NetMeeting and a TV monitor showing the video. Other technologies used were *ICEMaker* (LSMD), which provided the linking between workstations and the shared spreadsheet. It enabled the members of the Design Team to publish specifications and parameters relevant to a particular subsystem. Voice conferencing software managed small group sessions by providing a single point of access for sharing multiple voice streams by telephone.

We examined several sets of data. We focused mainly on a space mission design that lasted nine hours over a week-long period, three hours on three separate days. Three different researchers traveled to Sites 1–3 during this period to observe, videotape, and interview team members. We thus had videotapes to compare the different team perspectives. Individuals at Site 1 wore wireless microphones and their voices were recorded on separate audio channels. Fourteen semi-structured interviews were conducted with team members at Sites 2 and 3 ranging from 30 to 60 minutes each. In addition, we used data from two other design collaborations. In one of them, three members of the research team observed Site 1 who collaborated on a conceptual design for a real space mission with two other teams at distributed sites around the U.S. These sessions lasted 18 hours and interaction at Site 1 was videotaped. Video data from a third collaboration involved only Site 1, so that we could get a better understanding of the team process and how it differed when the team collaborated across distance.

The videotapes from the nine-hour session were transcribed and coded. The coding scheme was identified through an iterative process of watching videotapes and discussing interaction. The design process was also parsed into different aspects based on field observations and a review of the design literature.

Coding was done for each actor's conversation turn in the transcript. Short utterances were not coded. We coded for factors affecting gaps and alignment in groups: human-human communication, technology used, the role of external representations, and organizational factors. We also coded design processes relating to design requirements, mutual construction of methods and terms, design rationale, information loss, and design coordination. We used grounded theory (Strauss and Corbin, 1998) to relate concepts.

Narrowing and extending the "space between" groups

Design-oriented discussion is inherently wide-ranging, complex and iterative, and not an orderly progression from initial idea to final design. The overall design process was semi-structured around a system of networked spreadsheets (shown on public displays) where design information was recorded. Though the teams used the spreadsheet to guide the topics and order in which they would be addressed, the Design Team often departed from this "agenda" when an issue emerged that evoked more in-depth discussion. When the issue had been resolved, or deferred pending future discussion, the Design Team returned to the spreadsheets to guide them to the next topic of discussion.

Different group methodologies used in design

Articulation is a necessary component to close the gap of the "space between" interacting groups. Team differences, however, can extend quite deep. Inter-organizational teams not only use different concepts and terms, but unique methodologies and processes for designing that are not readily visible to other teams. In collaborative design, often just the "point result" of a calculation is communicated and other teams are not aware of the methodology used to achieve the result. Thus, although results and terms are visible during the interaction, the deeper processes and methodologies used to create results can be invisible to other teams. They must be explicitly communicated.

Figure 1 illustrates how different teams in the same engineering disciplines employed different methodologies involving a fairly "standard" computation. Two sites had proposed different values for a mass contingency factor, each with a different rationale (Site 1 used guidelines and Site 2 used a subsystem model), and each led to different consequences for mission duration. The line numbers in

Figure 1 indicate the order in which the various points are made/proposed, demonstrating that the order of their emergence in the design process is clearly different than their relationships, in terms of rationales and consequences.

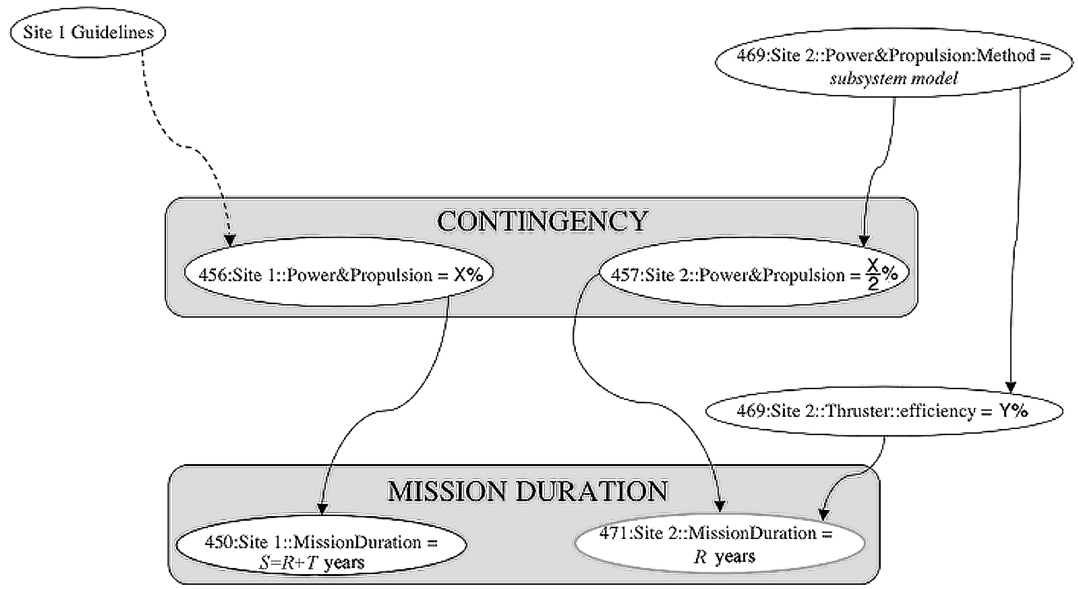

Figure 1. Different methodologies employed by two teams for the same design concept.

This figure illustrates a small part of a broad design space which can be further partitioned into two domains, defined by who actually makes the choices. Choices made by the Design Team are part of the "internal" design space; otherwise it is "external" made for example by the physics of a situation, (e.g. Mars cannot be moved closer to Earth to shorten the trip time), requirements imposed on the design space by external organizations (e.g. the customer), or by discipline-specific best practices. Internal choices are subject to change by the design team; external choices are not. In Figure 1, the "Site 1 Guidelines" used to support a factor of *X%* for contingent mass in the power and propulsion system is an example of a requirement from the external design space that can only be over-ridden with difficulty (such as with a waiver) or not at all (such as might be expected for human safety requirements).

The multidisciplinary aspect of the Design Team's process is not emphasized in the simple illustration in Figure 1. To reconcile the different approaches, in order to converge to a mutually-acceptable solution, each site must articulate their own methodologies and rationales and adopt the common solution.

We discuss four situations that illuminate how articulation is done across sites. In each case, different approaches to the same task emerged that reflected differing interpretations by the unique social worlds. As each site learned of the others' approaches, the sites constructed new methods, with new terminology, to apply in an emergent hybrid solution that satisfied each of their different needs.

However, although hybrid terms were agreed upon, the deeper methodologies and processes failed to be adopted, which led to consequences.

Contingent mass: hybrid methods that are not accepted

Though two successive hybrid solutions cooperatively emerged to reconcile different spacecraft masses, neither solution was accepted by the Design Team.

Key drivers of any space mission design are the masses of the various components of the technological system. These masses are initially estimated and iteratively refined as the design evolves. To allow for unanticipated growth of mass values at later stages of development, mass estimates are increased by a "contingency factor."

Sites 1 and 2 each had responsibility to design a trajectory for this mission (see next example) and to do so, needed to develop a mass estimate for the spacecraft, including "contingent mass." Site 1 employed a *top-down* approach by applying a single, default scaling factor of *X%* (defined by their site's design guidelines) to the entire spacecraft. Site 2 used a *bottom-up* approach by estimating contingent mass for each subsystem, as a percentage of that subsystem's mass, and then averaging the results over the entire craft. This resulted in a scaling factor of approximately *X/2%*. These highly discrepant values had significant impact on the mission duration and cost.

A sidebar, with representatives from Sites 1–3, met and converged on a hybrid approach that applied Site 1's default scaling factor to some of the subsystems (but not all) in Site 2's model and which yielded an intermediate value of approximately *2/3X%*. This also corresponded with an organization-wide default value. To determine to which subsystems different factors would apply, a new term "*validity*" was created based on prior experience with subsystems.

However, a facilitator at Site 1 continued to push for their default value, arguing that experience had shown them that *X%* was necessary to "sell" a proposal to those who make funding decisions. Articulation resulted in the emergence, and tentative adoption, of a second hybrid solution, in which Sites 1 and 2 would each use their own methods for design aspects for which they were responsible. At the end, the actual decision was deferred, by the Site 1 facilitator, until a sidebar with other facilitators could re-consider the issue. Site 2 was surprised by this resistance, considering the common agreement during the sidebar, and by Site 1's conservative approach. Site 2 expressed apprehension that such a policy-based approach would defeat their more mathematical approach.

Thus, in spite of considerable effort to cooperatively resolve such a key driver, articulation failed to deliver a solution, perpetuating the "space between" sites.

Trajectory design: decontextualized values that are not accepted

One of the most important consequences of mass determination is the range of viable trajectories that result. In this example, the sites' unique methodologies resulted in different trajectories that impacted the duration of the mission. Articulation revealed that sites used differing definitions of "trajectory." This led to the emergence of new terminology but it failed to be accepted.

The importance of the trajectory in the design was to accomplish the mission goals in a window of time bracketed, in the near term, by programmatic limits on how quickly the spacecraft could be built and, in the far term, by changing physical conditions at the destination that would impede mission goals. In general, a deep space trajectory consists of several distinct stages: launch from Earth to orbit, spiraling out from this orbit, transit to the destination, spiraling into orbit at the destination, and then performing operations in orbit to achieve the mission's goals. The time to complete all stages determines the length of the mission, which must fit into the specified window of time.

Sites 1 and 2 had primary and secondary responsibilities for the design of a trajectory for this mission, which had an unusual mass/power-ratio. Each site had different definitions of trajectory. In a top down manner, Site 1 began with the mass of a previous mission design and linearly scaling up from that mission's mass/power-ratio, it proposed a "*flight time*" of *S* years for a trajectory. In contrast, using a bottom-up approach, Site 2 developed a mass estimate for each subsystem that yielded a "*trip time*" of *R* years. Site 1's proposal combined the spiral-out and transit stages; Site 2's estimate also included the spiral-in stage.

Following Site 2's proposal, Site 1 adopted Site 2's term "*trip time*," though no value changed to indicate an expanded definition of "trajectory." No decision was immediately reached between the two proposals, as the discussion identified "contingent mass" as the critical distinction between the two proposals and this issue was deferred to a sidebar held between Days 1 and 2.

Towards the end of Day 1, the operational stage was included with flight/trip times and the hybrid term "*mission duration*" was inaugurated by Site 1. Site 2, however, never adopted this new hybrid term, preferring to use their term "*trip time*" to include this additional, operational stage. Later on Day 1, Site 1 reverted back to its term "*flight time*" to characterize its trajectory proposal. Thus, the hybrid term was never fully adopted by the Design Team.

These difficulties in articulating the different terms were increased by similarities in their related numerical values during subsequent discussion. In addition to the "*S* year" and "*R* year" values introduced, some discussion included the operational stage (required to be *T* years to achieve the science goals, where $S=R+T$). This led to some confusion at Site 1, as different people interpreted "*S* years" as referring to both the initial flight time and to the expanded "mission duration" ($R+T$ years). In the articulation, the confusion over the decontextualized values surfaced and agreement on common terminology failed to be adopted.

Power mode: new terminology that is not adopted

The power available to a spacecraft is also a key driver of space mission design. In this example, new terminology was developed to reconcile different power mode views, yet it was not adopted.

Power is a critical spacecraft subsystem. The provision of power can interfere with the ambient electromagnetic environment that the science instruments are measuring. It can also interfere with the operational instrumentation supporting navigation, telecommunications, and command and control of the craft. To work around this, different power modes are considered for different stages of the mission, to interleave the operational and science activities. These modes range from “no power” (0%) to “full power” (100%). Full power is usually defined as only a little more than is necessary to do the job because superfluous power generally has penalties of increased mass and higher costs. A 0% power mode has risks of loss of control and the freezing of critical systems. In this mission design, Sites 2–4 had the primary and secondary responsibilities for various aspects of the design’s power system, providing expertise that Site 1 lacked.

The initial discussion of power modes, initiated by Site 1, used the term *“power down”* to refer to a 0% power mode. When discussion concerned engineering safety, led by the Site 3 facilitator mid-Day 1, the term *“shut down”* was used to refer to this same 0% power mode. When this discussion led to other aspects, at the end of Day 1 by Site 2, the preferred term reverted to *“power down.”* This pattern of using two different terms to refer to the same design concept, in the context of engineering safety and in other contexts, was generally repeated on Day 2 when Site 2 reported on a sidebar that used the two terms according to their different contexts.

The two contexts had different impacts on the overall mission. There was no science requirement for a 0% power mode, but engineering safety did require one. Midway through Day 2, a hybrid solution emerged in which there would be no 0% power mode planned during science operations, but such a mode would be planned for all other phases of the mission. In addition to the persistent use of two different terms to refer to the same concept, a facilitator at Site 1 also introduced a third term, “*safing*,” to refer to the same 0% power mode in the engineering safety context. Failure to converge on specific terminology at least reflected, and perhaps contributed to, the failure of the team to adopt this hybrid solution.

These multiple terms persisted through Day 3 and, to complicate matters further, Site 4 proposed new terms to refer to five distinct power modes ranging from “no power” to “full power”. At the end of this discussion, at nearly the end of the overall design session, the team finally converged on the definitions of two terms, although not in time to incorporate them into other aspects of the design developed earlier in the session. “*Shut down*” was defined as the 0% power mode and “*power down*” was defined as referring to a value between 0% and 100%.

Multiple terms, representing similar aspects of design, increase the opportunity for confusion. Though attempts to construct consistent terminology were articulated, the entire Design Team did not consistently adopt these terms.

Rotational deviation margin: neglecting to address deeper rationales

This example shows how different sites, using distinct rationales, specified conflicting values of a significant design parameter.

Rotational Deviation (RD) results from the dynamics of the engine (produced by Site 4) and from spacecraft factors. Each RD (of the engine, spacecraft, or both) can be expressed by either its raw value, or a padded value (the raw value multiplied by a precautionary margin called RDM). RD is a significant design value because it negatively impacts the precision of on-board instruments (designed by Site 1). Therefore, the Design Team must specify tolerable values of RD and adopt a value of RDM based on organizational safety requirements and historical standards, to design the engine and instruments accordingly.

On Day 2, Sites 1 and 4 initially converged on an RD value of 5°. The Site 1 facilitator asked about the spacecraft RD and its RDM. Coincidentally, this RD also happened to be 5° with RDM of 2. A question from Site 4 revealed that team members were using the number 5° liberally without explicit reference to whether it was the raw engine RD or the padded spacecraft RD, both equal to 5°. Later, both sites referred to total RD. It was clear to the observers that "total" for Site 1 meant both RDs, whereas "total" for the Site 4 participant meant engine RD, for which he was responsible. Each site was thinking about its own component.

The team then focused on specifying protection for instruments. This revealed a difference of design methodology between sites. The Site 1 methodology began with 5°, multiplied it by RDM, and designed instruments to withstand 10°. The Site 4 methodology instead began with 5°, divided it by the margin, and designed the engine for the resulting number, namely 2.5°, while expecting the instruments to withstand 10°. This suggested that the Site 4 participant still did not use the term "*margin*" with the same understanding as that of Site 1.

The customer then asked Site 4 to explore a more desirable engine RD, namely, by cutting the value in half so that the total RD is 5°. Exchanges, clarifications, and repetitions of numbers ensued, further complicated by the fact that there was a different significance to the factor 2, and to the value 5°. So when someone used the number 5°, without explicit clarification (which was frequently done), it could have meant raw engine RD, desirable padded engine RD, padded spacecraft RD, or desirable raw total (engine and spacecraft) RD.

In conclusion, the team recorded the decision to use engine RD as 5° with no RDM, and spacecraft RD as 5° with RDM of 2. This decision did not explore the deeper rationales for selecting this choice of numbers. Though the numbers were settled on, the deeper methodologies underlying how the numbers were generated were still not resolved.

Surface and deep articulation

These four cases are complex (space mission design is very complex) but they all illustrate the same idea: hybrid terms and even simple processes were agreed on, but the deeper differences of the different team approaches were not sufficiently articulated. When an issue came to the attention of the Design Team, to some extent each site supported their proposed solutions—their design rationales—by explaining on the public channel "where they were coming from." New terminology was constructed to distinguish it from site-specific terminology. The hybrid solutions were accepted by the Design Team, but in language only. Though the facilitators initially believed that a hybrid solution was achieved and would therefore be adopted, in fact it was not.

In group-to-group collaboration, a "group-centric" view of the design can occur. The processes and methodologies at one's own site are visible and known. Articulation is triggered when results—in the Design Team's case, calculations or point estimates—are incommensurate, or seem questionable to another site. The deeper processes and methodologies used to calculate those results are not visible. Though the Design Team successfully created new hybrid terms, they neglected to reconcile the deeper methodologies and processes that lay "underneath" the terms. Although on the surface the term was agreed upon (at least initially), the many layers of methodologies and meaning that the term represented were not similarly reconstructed for the entire Design Team.

Insufficient translation is a problem endemic to different specialist groups because of their development of different languages, cultures, and approaches (Heath and Staudenmayer, 2000). It is not surprising that teams approach design differently (though they are all in the aerospace engineering domain in the same organization[1]) because through long term team relationships, methodologies and processes become embedded within team cultures. Even though hybrid solutions were formed to achieve consensus, translation of the deeper processes did not occur. As a result, the "space between" groups remains wide leaving much potential for error in design.

Information bridges across sites

We propose that in large networks people can be information bridges that affect the "space between" distributed teams. These information bridges may not only function to pass on content from the local site, but they may also represent the site in articulation. This is exactly what we found.

The VTC was a major part of the synchronous collaboration though information was also communicated through sidebars with telephone and by entering results into publicly displayed spreadsheets. The data suggested to us that

[1] Site 4 was in a different organization though also in the engineering domain.

the VTC was important to focus on for information transmission. In all four of the cases described earlier, most discrepancies between methods were discovered through conversation over the VTC channel. We should note that discrepant values were also identified in the shared spreadsheet.

Group-to-group collaboration technologies enable anyone at any site to freely participate. For example, anyone can speak through the VTC or enter their results into the networked spreadsheet that is publicly displayed. On the other hand, if everyone speaks freely without social protocols, chaos can possibly ensue. We were interested in exploring the social mechanism by which information was conveyed across sites using the VTC channel. In particular, we were interested in the role of information bridges in communicating information.

To explore this, we transcribed the VTC communication into conversational "exchanges." Such exchanges are contiguous streams of one speaker's speech and should not be confused with conversational "turns." In this transcript, a classic conversational turn may be split over several exchanges as people speak over and interrupt one another. We felt this approach was closer to the actual conversation and might afford opportunities to explore the impact of such conversational missteps on the design process. We coded each conversational exchange to determine whether it involved articulation or design content[2]. Using Gerson and Star's (1986) definition of articulation, we coded for any exchange that concerned "*all the tasks needed to coordinate a particular task, including scheduling subtasks, recovering from errors, and assembling resources*" (p. 258). This also included any discussions to reconcile different perspectives. We also coded for any exchange in which design content was discussed. Extraneous discussion not fitting into these other categories (e.g. jokes) was coded as "other."

Figure 2 shows the proportion of time the Design Team spent in articulation. There are several interesting aspects of this data. First, note that articulation occurred fairly regularly over the entire design session. One exception shows articulation increasing in frequency at the beginning of Day 3, when the designers know they have a limited time left. By correlating this with the transcript, we propose that the articulation provides a frame for the next design activity. The Team articulates their plans, designs in depth, and then coordinates to "recover" from the depth. The Team then moves on to the next topic.

We found that people whose roles involved coordinating the design process at their site tended to act as "bridges," speaking as representatives of their site to the entire Design Team. This is not a typical role for facilitators whose primary responsibility was to keep the team at their own site on track. The facilitators evolved their roles into becoming information bridges across sites.

---

[2] One member of the research team was a domain expert in space mission design.

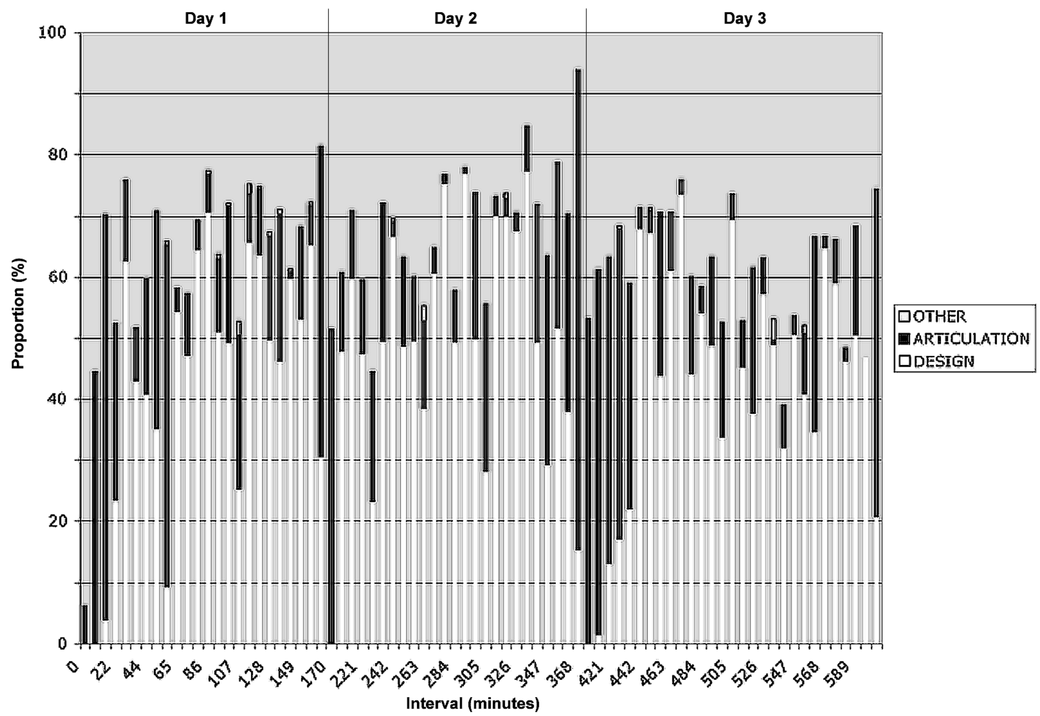

Figure 2. Proportion of time spent in articulation and design activities over the entire session.

The engineers who were more directly engaged in designing spoke less. Facilitators accounted for 61% of the public conversation (Table 1). Site 1's dominance of the common communication channel is also reflected, accounting for 70% of the shared conversation. Moreover, after agreeing on using voice protocols to identify themselves, Site 1 provided only 13% of self-identifications over the entire three-day period. Site 2 provided 58% of the self-identifications and Site 3 provided 29% though it contributed only 12% of the public conversation. This suggests that Site 1 did not make the extra effort to adapt their normal procedures to this new Design Team environment, as the other sites had done.

	Site 1	Site 2	Site 3	Site 4	TOTAL
Facilitators	45%	7%	9%	n/a	61%
Other team members	25%	6%	3%	5%	39%
TOTAL	70%	13%	12%	5%	100%

Table 1. Proportions of VTC-communicated content.

Second, we examined how much of the time was spent on articulation activities versus working on design content (e.g. calculations, meeting requirements). We discovered that about 70% of the discussion was spent discussing design content, and about 30% was spent on articulation (Table 2). Thirty percent is perhaps not so high considering the great amount of articulation needed, as the four earlier examples illustrated. Furthermore, the information bridges spent three times as much of their discussion time on articulation discussion (21.4%) as the rest of the team members, who spent only 7.0%. Thus, it appears that information bridges represented their local team in articulation. They also spent slightly more time on design discussion (39.4%) compared to the other team members (31.7%) which suggests that they also served as bridges for their teams in public discussion about design.

Thus, in this group-to-group collaborative setting, most of the discussion over the public VTC channel was done by "information bridges" representing their sites. To what extent can this result be generalized to other group-to-group collaborative settings? It seems reasonable that most information would pass publicly across sites through information bridges simply due to the large number of actors at multiple sites. The implications for design are that much of the articulation is dependent on these bridges. Not only is the quality and depth of articulation governed by such bridges but also the amount of articulation. Even if the facilitators agree on a hybrid solution (as in the four cases in the last section), this is far different than reconciling team methodologies. The fact that the rest of the team engaged in public articulation only 7% of the time may have hindered the Design Team from converging toward common methodologies and processes.

Conversation	**Facilitators**	**Other team members**	**TOTAL**
Design	39.4%	31.7%	70.5%
Articulation	21.4%	7.0%	28.4%
Other discussion	.2%	.3%	.5%

Table 2. Proportion of time spent on design and coordination/articulation conversation over the three-day mission design.

Lost information within and across sites

In the last section we discussed "public articulation," i.e. what occurs over the public VTC channel, that is available to the entire team. In this section we report how sidebars (subgroup conversations) are an important part of the articulation process in group-to-group distance collaboration. In the cases described earlier, in articulating mass contingency, a facilitator sidebar was formed to reconcile the approaches. Similarly, a sidebar was formed to articulate the design of power modes. It is critical in the articulation process that the results of the sidebar be communicated back to the team. Yet we discovered this was not the case. When critical information from sidebars are lost to the larger team, this further expands the "space between" distributed groups.

During the design, many sidebars occurred in parallel to the main VTC channel discussion. These sidebars occurred to gather information or to work on a small segment of the design, but mostly they convened for articulation purposes, i.e. to reconcile different team perspectives. In short, if a problem could be solved by relegating it to a sidebar, then it was. Sidebars were formed either by the team members, or they were assigned by a facilitator. They occurred both within a site and across sites, using telephones and the voice conferencing technology.

In fact, the different teams entered the collaboration with different practices of conducting articulation. It was normal practice at Site 1 to defer to a sidebar any topic in disagreement or requiring specialized expertise. These sidebars involved either the parties in disagreement or expertise from multiple sites and/or disciplines. Our data confirm how multiple perspectives are reconciled in sidebars to yield a commonly accepted solution. In the earlier contingent mass example, there was strong resistance from Sites 2 and 3 (noted in two instances) to defer topics to sidebars when there was disagreement among the sites. As the Site 3 facilitator expressed: "*if we don't have consensus on it, I think it should be brought out.*" These sites expected contentious issues to be resolved publicly, instead of being deferred to sidebar discussions.

Coding the sidebars over the three-day session revealed that the coordination of all sidebar discussions fell into four categories: suggestions to convene a sidebar discussion and actual set-up, public reference to sidebars (after their suggestion and before their reports), the report of results, and the resolution of technical difficulties[3]. The public references to sidebars suggests that the entire Design Team was made aware that a sidebar was in progress, and with it, the expectation that the results would be communicated back to the team. Thus, the Design Team knew that people in a sidebar were articulating an issue.

	Site 1	Site 2	Site 3	Site 4	TOTAL
Suggestion/Setup	50	2	3	3	58
Reference	28	6	4	1	39
Report	4	2	1	0	7
Resolution	4	1	1	0	6
TOTAL	86	11	9	4	110

Table 3. Coded numbers of specific phases of sidebar processes, for 22 distinct topics, identified by Site.

The most surprising observation is the asymmetry between the number of references suggesting to convene and set up sidebar discussions and the relatively few reports back to the entire Team (Table 3). Though 58 sidebars were set up, in only seven cases (6%) were they publicly reported back to the team. Even though

[3] Codes included both collocated and distributed sidebars.

it is likely that the results of the sidebar discussions were implicitly communicated back to those interested, e.g. via changes made to the shared spreadsheets, their public call for, and subsequent public coordination, created an *expectation* that the information would be forthcoming via a public channel as well. Sites 2 and 3 made at least four references to such expectations. Thus, when sidebar results, especially those concerning articulation, are reported back to either one site only or not reported at all, then this is a way for common meaning to be lost in group-to-group collaboration.

Discussion

Designing for group-to-group distance collaboration presents a great challenge to CSCW. Currently the development of new room-size technologies that take advantage of large bandwidth is proceeding at an astounding pace, as with Access Grid Technologies. In this study we have attempted to understand problems in working with this new type of collaborative configuration. The extent to which technologies, communication, translation and articulation can close the gap of the "space between" is a challenge for distributed large-scale group-to-group meetings to be able to engage in meaningful and effective collaboration.

First, it is difficult for teams to achieve common aligned methods. Bechky (forthcoming) describes that knowledge is contextual in different organizational communities. Decontextualization occurs as different groups use different words and concepts to refer to the same concept. Recontextualization occurs when individuals use methods (e.g. providing a tangible definition) to arrive at a common understanding. We found that the team differences extended deeper than different terms/concepts. Design methodologies and processes were developed in different organizational contexts resulting in deep differences between the teams.

These methodologies are not readily visible to other teams across distance. Though hybrid solutions were achieved as a result of articulation, the local teams failed to adopt them. Team processes are deeply embedded (e.g. they may be local site design guidelines) and may not readily change even though a common solution is nominally agreed upon. Group-centric views, found in intergroup collaboration (Ancona and Caldwell, 1992), can inhibit the adoption of methods from outside the group. Early on, Gerson and Star (1986) proposed that a conceptual basis for the design of computer systems must be based on an understanding of articulation. The design of large-scale group-to-group collaborative technologies must enable not only point results but also the *visibility of team methodologies and processes* so that they can be articulated. For example, algorithms and formulae used by teams could be easily linked to the point result in the shared spreadsheet so that they are visible to the entire distributed set of teams. Increasing visibility is one small step, however. Adopting new methodologies is another large leap. Research needs to focus on how different teams can overcome group-centric views to adopt methodologies common to a larger-scale collaboration.

The use of information bridges was a major part of this design effort. We expect that such bridges are common in group-to-group collaboration efforts. A technology requirement is for a tool that can track team progress at a local site to enable the person serving as an information bridge to better represent the team. One possibility is to display a spreadsheet publicly to all sites that presents each sites' progress. This enables multiple "eyes" to detect discrepancies and nonalignment. This also can trigger opportunities for articulation.

We discovered that results of small group discussions that concern articulation (and other topics) rarely get reported publicly and integrated back into the larger distributed team. An important value of publicly conveying the results is that it documents the design rationale which can be later accessed. No mechanism existed in the collaborative technology for sharing the results of sidebar discussions with the entire team. A requirement for large-scale group-to-group systems is to provide a mechanism for tracking sidebars, their topics, and for channeling their results back to the larger team.

Technology can serve to close the gap of the "space between" actors and groups, but it can also function to widen the gap. We found examples with the sampling approach of the VTC and with the networked spreadsheets on the public display. As described earlier, to maximize bandwidth utilization, the video-teleconferencing service switches each site's local display between the remote sites, depending upon their most recently sampled volumes. This approach enabled people at a local site to get a video image of who was speaking currently or recently. But this has two drawbacks for supporting multiple site collaboration. First, during the volume sampling period, viewers see the site that was most vocally active before the sampling period, not the site currently speaking. This discrepancy was noticed at Site 2 and mistakenly identified as both a problem of temporal synchronization between the audio and video streams and as an effect of transmission delays.

The second drawback is more subtle. As designed, the system should work as intended if all sites vocalize in equal amounts and volumes. However, if one site dominates the discussion (e.g. Site 1), the video stream may tend to reinforce that dominance relationship. Since the VTC system will not display the dominant site to itself, viewers at the "speaking" site will primarily see video streams of people from the other sites listening to them (more or less equally) and not talking. Conversely, the less vocal sites will be presented with video streams primarily from the dominant site. Some research suggests that people evaluate persons in visual images who speak much more than listen as being "more in control" than if persons in the image seem to be primarily listening (Dovidio and Ellyson 1982). Conversely, visual imagery of a person listening more than speaking will be evaluated as "less in control." We suggest that the VTC sampling could have reinforced the verbal dominance of Site 1 by making them seem more in control to the other sites while making the other sites seem less in control to Site 1.

Conclusion

To the extent that the "space between" remains wide between entire teams at a distance, the risk for errors increases. With time, congruent agreed-upon practices and methodologies could likely be reached to narrow the "space between" collaborating groups. Currently ad hoc collaborations appear to be common in group-to-group collaborative situations, i.e. when groups convene across distance for specific discussions. Long-standing group-to-group collaborations are still rare, though as technology improves they may become more common. Providing large interfaces, e.g. to display video images or large data sets, are a first step, but may provide only a "quick fix" to enable such large-scale collaboration. An understanding of the adaptation and alignment of local team practices in different organizational contexts is not keeping pace with the development of group-to-group collaborative technologies.

Acknowledgements

We would like to thank Rebecca Wheeler, Paul DeFlorio, Steve Prusha, and members of the Design Team. This material is based upon work supported by JPL/NASA Grant No. 1240133 and the National Science Foundation Grant No. 0093496.

References

Access Grid (2002): http://www-fp.mcs.anl.gov/fl/accessgrid/default.htm

Ackerman, M. S., Hindus, D., Mainwaring, S. D., and Starr, B. (1997): 'Hanging on the wire: A field study of an audio-only media space', *ACM Trans. on Computer-Human Interaction*, vol. 4, no. 1, pp. 39-66.

Ancona, D. G. and Caldwell, D. F. (1992): 'Bridging the boundary: External activity and performance in organizational teams', *Admin. Science Quarterly*, vol. 37, no. 4, pp. 634-665.

Bavelas, A. (1950): 'Communication patterns in task-oriented groups', *Journal of the Acoustical Society of America*, vol. 22, pp. 271-282.

Bechky, B. (forthcoming): 'Sharing meaning across occupational communities: The transformation of understanding on a production floor', *Organization Science*.

Bradbury, H. and Lichtenstein, B. (2000): 'Relationality in organizational research: Exploring the space between', *Organization Science*, vol. 11, pp. 551-564.

Bradner, E., Kellogg, W., and Erickson, T. (1999): 'The adoption and use of 'Babble': A field study of chat in the workplace', *Proceedings of the ECSCW'99*, Copenhagen, pp. 139-158.

Carroll, J. M. (2000): *Making Use: Scenario-based Design of Human-Computer Interactions*, The MIT Press, Cambridge, MA.

Dovidio, J. F. and Ellyson, S. L. (1982): 'Decoding Visual Dominance: Attributions of Power based on Relative Percentages of Looking-while-Speaking and Looking-while-Listening', *Social Psychology Quarterly*, vol. 45, no. 2, pp. 106-113.

Finholt, T.A. (2003): 'Collaboratories as a new form of scientific organization', *Economics of Innovation and New Technologies*, vol. 12, pp. 5-25.

Fitzpatrick, G., Kaplan, S., and Mansfield, T. (1996): 'Physical spaces, virtual places and social worlds: A study of work in the virtual', *Proceedings of CSCW'96*, pp. 334-343.

Freeman, Linton C. (1991): 'Centrality in valued graphs: A measure of betweenness based on network flow', *Social Networks*, vol. 13, pp. 141-154.

Gerson, E. M. and Star, S. L. (1986): 'Analyzing due process in the workplace', *ACM Transactions on Office Information Systems*, vol. 4, no. 3, pp. 257-270.

Heath, C. and Staudenmayer, N. (2000): 'Coordination neglect: How lay theories of organizing complication coordination in organizations, in B. M. Stay and R. I. Sutton (eds.): *Research in Organizational Behavior*, Elsevier, Oxford: pp. 153-191.

Hutchins, E. (2000): *Cognition in the Wild*, The MIT Press, Cambridge, MA.

LSMD (Laboratory for Spacecraft and Mission Design): ICEMaker, http://www.lsmd.caltech.edu/tools/icemaker/icemaker.php

Mark, G. and DeFlorio, P. (2001): 'An experiment using life-size HDTV', *Proceedings of the IEEE Workshop on Advanced Collaborative Environments*, San Francisco, CA.

Mark, G., Grudin, J. and Poltrock, S. (1999): 'Meeting at the desktop: An empirical study of virtually collocated teams', *Proceedings of ECSCW'99*, Copenhagen, pp. 159-178.

Nardi, B., Whittaker, S., and Bradner, E. (2000): 'Interaction and outeraction: Instant messaging in action', *Proceedings of CSCW2000*, pp. 79-88.

Nardi, B., (ed.) (1996): *Context and Consciousness: Activity Theory and Human-Computer Interaction*, The MIT Press, Cambridge, MA.

Olson, G. M., & Olson, J. S. (2000): 'Distance Matters', *Human-Computer Interaction*, vol. 15, no. 2/3, pp. 139-178.

Roberts, K. and Yu, K. (2003). 'The space between: Partitioning and component focus as sources of error in healthcare.' *Unpublished manuscript.*

Robertson, T. (1997): 'Cooperative work and lived cognition: A taxonomy of embodied actions', *Proceedings of ECSCW'97*, Lancaster, UK, pp. 205-220.

Ruhleder, K. and Jordan, B. (1999): 'Meaning-making across remote sites: How delays in transmission affect interaction', *Proceedings ECSCW'99*, Copenhagen, pp. 411-429.

Schmidt, K. and Bannon, L. (1992): 'Taking CSCW Seriously: Supporting Articulation Work', *Computer Supported Cooperative Work*, vol. 1, no. 1-2, pp. 7-40.

Strauss, A. L. and Corbin, J. M. (1998): *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*, Sage Publications, Thousand Oaks, CA.

Vaughan, D. (1997): *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*, University of Chicago Press, Chicago, IL.

Weick, K. E. and Sutcliffe, K. M. (2001): *Managing the Unexpected: Assuring High Performance in an Age of Complexity*, Jossey-Bass, San Francisco, CA.