

MODELLING DISTRIBUTED KNOWLEDGE PROCESSES IN NEXT GENERATION MULTIDISCIPLINARY ALLIANCES^{*}

Alaina G. Kanfer
National Center for
Supercomputing Applications,
University of Illinois at Urbana-
Champaign (UIUC)
alaina@ncsa.uiuc.edu

Bertram C. Bruce
University of Illinois at Urbana-
Champaign (UIUC)
chip@uiuc.edu

Caroline Haythornthwaite
University of Illinois at Urbana-
Champaign (UIUC)
haythorn@uiuc.edu

Nicholas Burbules
University of Illinois at Urbana-
Champaign (UIUC)
burbules@uiuc.edu

James Wade
University of Wisconsin at
Madison
jwade@bus.wisc.edu

Geoffrey C. Bowker
University of California at San
Diego
bowker@ucsd.edu

Joseph Porac
Emory University
joe_porac@bus.emory.edu

^{*} These ideas have been developed within the Distributed Knowledge Research Collaborative with valuable input from Joyce Brown, Mike DeVaughn, Michelle Guizac, Vikas Khajanchi, Karen Lunsford, Tim McDonough and Katie Vann. An earlier version of this paper received the *AIWORC-Bell Atlantic Best Paper* award at the Academia-Industry Working Conference on Research Challenges (AIWoRC'00) in Buffalo, NY, April 27 - 29, 2000. This research was funded by a grant from the U.S. National Science Foundation, Knowledge and Distributed Intelligence program, <see <http://www.ncsa.uiuc.edu/edu/dk>>.

ABSTRACT

Current research on distributed knowledge processes suggests a critical conflict between knowledge processes in groups and the technologies built to support them. The conflict centers on observations that authentic and efficient knowledge creation and sharing is deeply embedded in an interpersonal face to face context, but that technologies to support distributed knowledge processes rely on the assumption that knowledge can be made mobile outside these specific contexts. This conflict is of growing national importance as work patterns change from same site to separate site collaboration, and millions of government and industrial dollars are invested in establishing academic-industry alliances and building infrastructures to support distributed collaboration and knowledge.

In this paper we describe our multi-method approach for studying the tension between embedded and mobile knowledge in a project funded by the U.S. National Science Foundation's program on Knowledge & Distributed Intelligence. This project examines knowledge processes and technology in distributed, multidisciplinary scientific teams in the National Computational Science Alliance (Alliance), a prototypical next generation enterprise. First we review evidence for the tension between embedded and mobile knowledge in several research literatures. Then we present our three-factor conceptualization that considers how the interrelationships among characteristics of the knowledge shared, group context, and communications technology contribute to the tension between embedded and

mobile knowledge. Based on this conceptualization we suggest that this dichotomy does not fully explain distributed multidisciplinary knowledge processes. Therefore we propose some alternate models of how knowledge is shared. We briefly introduce the setting in which we are studying distributed knowledge processes and finally, we describe the data collection methods and the current status of the project.

INTRODUCTION

We are rapidly moving toward a world in which knowledge is constructed collaboratively at a distance by multidisciplinary teams, supported with an electronic communications and information infrastructure. The exponential growth of knowledge [38] has made it nearly impossible for any organization to exist in isolation. Thus, the networked organization, or alliance, is an increasingly common structural form within and between science, government, business, and non-profit organizations [17,32].

An alliance is a collection of organizations that have entered into collaborative relationships usually involving multiple channels of communication and knowledge diffusion across disciplinary or organizational boundaries. The proliferation of alliances has been stimulated, in part, by emerging electronic technologies that support collaboration, ranging from email and intranets to desktop videoconferencing and collaborative data mining and visualization. In fact, across business and government organizations, tremendous resources and significant financial investment have been devoted toward these efforts.

Alliance organizations depend on effective virtual collaboration. However, empirical research exploring collaborative knowledge processes and the function of electronic infrastructures in distributed work groups reveals an important contradiction. This contradiction emerges, on the one hand, from observations that authentic knowledge processes are somehow *embedded* within specific practices and interpersonal exchanges [30,7,31 & 37]. On the other hand, successful use of electronic infrastructures to support knowledge processes among distributed teams depends on knowledge being made *mobile*, that is, transferable across people located in different places. Thus it may be very difficult to transfer authentic embedded knowledge beyond the specific setting and set of people in which the context for that knowledge is consensually understood. This is the contradiction we intend to investigate in our research: that is, *the nature of knowledge processes in groups and the goals of electronic infrastructures to support distributed knowledge processes may be in direct conflict with one another*. The resolution of this conflict may even be so basic as to involve rethinking our understanding of knowledge processes.

Thus, we are presented with a situation where government and industry are investing millions of dollars to develop infrastructures to support knowledge-based alliances that are based upon conflicting principles. We believe it is critical to understand this conflict and sort out when and how various knowledge processes can be distributed. What better context to learn about electronically supported distributed knowledge processes than in the multi-million dollar national experiment, Partnerships for Advanced Computational Infrastructure (PACI), funded by the National Science Foundation? This national initiative to prototype the infrastructure for the 21st century is based on an alliance model and has available the most sophisticated forms of communication technologies.

We, as a distributed (at four sites across the US) and multidisciplinary team ourselves, are studying the construction and sharing of knowledge among multidisciplinary team members in one of the two PACI funded partnerships, the National Computational Science Alliance (the Alliance). The Alliance consists of over sixty educational, government and industry partner organizations, with the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign (UIUC) as the leading edge site. The six teams we are studying are known as the Application Technologies (AT) teams. Each team is involved in creating computational tools to support particular sectors of research important to academia and industry, and include members at multiple sites within the US. The teams represent innovative and leading edge collaborations across disciplines, often across disciplines that have little history of working together. For example, the Environmental Hydrology team brings together researchers in atmospheric sciences, hydraulic research, computer science, systems ecology, environmental engineering, meteorology, space and science engineering, geography and computational fluid dynamics to examine large environmental trends in water management.

Our investigation of these teams and their knowledge processes is multidisciplinary, multi-level, and multi-method. We bring together in our research team investigators from social science, computer science, history, business, psychology, philosophy, and information science to explore the knowledge processes that are involved in virtual collaborative work. We are interested in how electronic technologies are used to support this knowledge production and sharing and how they can be used in other alliances. This following section describes the theoretical underpinnings of our work and relevant previous research findings.

THEORETICAL AND EMPIRICAL UNDERPINNINGS

Our research is based upon a number of research streams that together suggest a widespread tension between mobile and embedded knowledge. For our purposes, knowledge can be considered "mobile" when it can be codified in a linguistic (written or oral) way and easily transferred or translated from one person or group to another. Knowledge is "embedded" in a social system when it is bound up with a set of communications, practices, and tools that make linguistic codifications difficult. The research literature in organization studies, education, sociology of scientific knowledge and computer-mediated communication illustrate the tension between rich knowledge that is embedded in interpersonal contexts, and the need to make knowledge mobile when it must be shared electronically among distributed team members [1, 8]. Our research focuses on how working scientists resolve this tension and the impact it has on their multidisciplinary collaboration and relationships with their home disciplines.

Some *organizational research* presents knowledge-based views of organizations suggesting that embedded knowledge is a key source of competitive advantage for organizations since such knowledge is hard to replicate and thus hard for competitors to imitate [27]. Organizations are thus encouraged to protect their proprietary knowledge by enhancing its embeddedness. In other words, partner firms in an alliance may try to protect embedded or tacit knowledge and make it more difficult to acquire [56]. On the other hand, the same research literature suggests that knowledge sharing (i.e. mobile knowledge) through interorganizational alliances plays an important role in the development of innovations because alliances permit a pooling of financial and intellectual resources that distributes the costs of large scale and expensive knowledge constructing projects [17]. Thus alliances require that certain degrees of knowledge be mobile in order for the alliance to be effective.

The issue of whether embedded or mobile knowledge better enhances knowledge construction also emerges in the *educational research* literature. The situated learning approach [31, 47], for instance, suggests that deep immersion in a setting facilitates learning, thereby implying that effective learning exploits embedded knowledge. In contrast, the value of asynchronous learning networks, a fundamental concern of the emerging field of computer-supported collaborative learning, is based upon the premise that embedded knowledge can be made mobile and sharable over various communications infrastructures [9, 28, 19].

As in the education literature, several perspectives on the *sociology of scientific knowledge* argue that the construction and sharing of scientific knowledge is largely dependent upon embedded interpersonal communications and collaborations [18, 21, 30, 35 & 36]. More and more scientific problems -- for example, the Human Genome

Initiative or studies of global climate change -- are very large-scale, resulting in scientific communities that are spatially distributed and often multidisciplinary in nature [6]. Thus in the process of trying to facilitate the collaboration necessary for multidisciplinary work, scientists often must transform knowledge that is embedded within highly specific domains into mobile knowledge that can cross several domains. There are indications that this effort to make embedded knowledge mobile and shared across multidisciplinary scientific teams results in a complex series of trade-offs between communication efficiency and preserving context. In particular, a 'pidgin' language [14, 10] or a 'boundary object' [42, 3] often must be constructed that is not scientifically precise but which gets the job done [11].

Similarly, the *social network* literature shows that closed groups only have access to information circulation among their members [4, 34] -- the knowledge is embedded, but not open to new influences. By contrast, groups whose members mingle with others can be exposed to new information -- information we argue may lead to new knowledge when embedded structures are amplified or embellished with information mobilized from other groups and brought into the focal group [15, 16 & 20].

The concern for making embedded knowledge mobile in an electronic infrastructure is not new to the study of *computer-mediated communication* (CMC). One of the primary concerns in this field is how to strike a balance between supporting work-oriented exchanges in an electronic infrastructure, versus socially-oriented exchanges which are essential to build trust and accomplish role formation, both of which are necessary for effective group processes [13, 33, 43 & 44]. In other words, our interpretation is that much of the CMC literature is concerned with how to create an electronic infrastructure that supports mobilizing embedded knowledge generally.

FACTORS CONTRIBUTING TO THE TENSION BETWEEN EMBEDDED AND MOBILE KNOWLEDGE

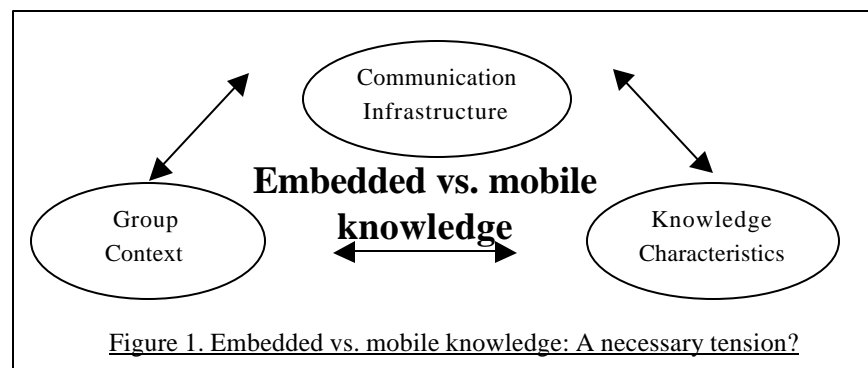
Conflicting results in the various research literatures described above undoubtedly have to do with the various contexts in which virtual knowledge processes are examined. While many factors affect the process of knowledge building within an alliance, we have identified, from the above literatures, three primary factors that interact, acting both as independent and dependent variables, to influence knowledge construction processes. These three factors,

which will provide a focus for our investigation, are: *characteristics of the knowledge* in a domain or multidisciplinary area; the *communications infrastructure*; and the existing *group context*.

The *characteristics of knowledge* include paradigmatic consensus, disciplinary span, boundedness and modification capacity. Each domain has its own ‘ecology of knowledge’ [7] that is recoverable through ethnographic analysis (in particular, participant observation at conferences and meetings and structured interviews), and content and bibliometric analyses of publications.

The *communications infrastructure* includes the communication media such as documents, electronic mail, teleconferences, videoconferences, teleimmersion, telephone, fax, letters, casual and scheduled face-to-face meetings, conferences, publications, Web sites, shared databases, events and transfer and exchange of group members. These media vary along several dimensions such as synchronous versus asynchronous, and permissive versus prescriptive [12]. Email can be an important form of scientific communication [5, 23]. Aspects of the communications use can be captured through ethnographic observation, transaction log analysis, structured interviews and questionnaires.

The *group context* includes the social structure when group members first meet, prior outcomes, publication frequencies and co-authorships, use of and attitudes toward technology, characteristics of the home institutions of group members, group demographics such as the disciplinary training of members, the position of members in their fields, and types and purposes of group interactions. Structured interviews, questionnaires and archival data can reveal characteristics of the group context.



As expressed in Figure 1, the impact of each of these variables on the others is realized as various aspects of embedded knowledge become mobile, and vice versa. Previous research has examined the group context and

communications infrastructure together with other characteristics such as productivity or satisfaction. Our approach contributes to the collective understanding of knowledge processes by extending the factors under consideration to include the technological infrastructure that supports the mobilization of knowledge, and the interaction this has with group context and knowledge characteristics.

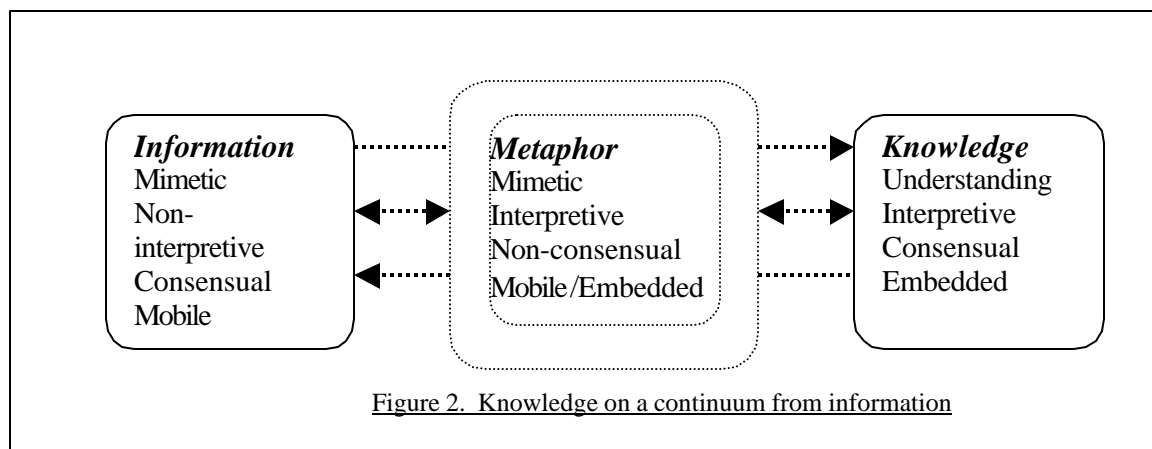
CONCEPTUALIZATION OF KNOWLEDGE

The resolution of the tension between embedded and mobile knowledge in multidisciplinary alliances may require us to think differently about knowledge. It is clear that in the analysis of knowledge mobility we cannot rely on a unitary theory of knowledge as a commodity that can be packaged in one place with a given context, transferred unchanged to a second site, and then readily inserted into this new context. Rather than considering a reified conception of “knowledge” and asking how it comes to be distributed, we expand our conceptualization of knowledge to include a continuum from information to knowledge, and from mobile to embedded.

Our view of knowledge to see it not as a final artifact, a set of understandings to be “disseminated” or transmitted to others, but as a set of conclusions generated out of, and continually adapted by, the social processes in which it is formulated, interpreted, and shared. We see these not as distinct stages, but as aspects of the knowledge process itself in any sort of group. Thus, we go beyond the easy information/knowledge dichotomy to explore the varying ways in which so-called “information” (which some might consider non-interpretive beliefs) is in fact continuous with “knowledge” (a highly interpreted and often tested set of beliefs) (See Figure 2).

Even the information and knowledge ends of this continuum are not completely distinct. The most apparently straightforward processes of information transfer can also involve very complex elements of interpretation, reading context, and adding inferences that go beyond the “information” itself. Conversely, even the richest and most critically reflective process of knowledge sharing involves elements of “imitation,” that take certain things for granted without examination.

One consequence of this is that in collaborative environments participants always have partly overlapping and partly distinct understandings of any matter under joint consideration. The line of understanding and misunderstanding is not dichotomous. Rather, in any collaborative process,



reinterpreting new information in light of one's assumptions and experiences is a necessary aspect of how any joint understanding happens. This also means that misunderstanding is a necessary, and in some ways a potentially beneficial, aspect of any collaborative situation. What is a 'misunderstanding' from one standpoint, may be a 'new twist' from another standpoint, and the active resolution of this misunderstanding may allow creative “abrasion” that leads to a new jointly created connectivity between standpoints or an understanding – new knowledge – about the problem. Thus, varying degrees of common interpretation and misunderstanding may in fact have very positive implications for knowledge processes in multidisciplinary distributed teams and may help us to resolve the tension between embedded and mobile knowledge.

MODELS OF KNOWLEDGE PROCESSES

Our question about the mobility of knowledge in alliances, then, focuses around what it is that is made mobile. What is codified in one discipline may not be accessible (mobile) to those in other fields because of the intellectual content and the amount of background needed to understand it. In order to *create* knowledge, rather than to *transfer* knowledge (as in the usual paradigm), what needs to be mobilized may not be the full knowledge of a discipline, but may be only a small part of the information necessary or sufficient to allow the creation process.

We suggest that such partial knowledge may be represented as a metaphor, and, indeed, metaphors have been identified as one modality through which scientific ideas get transmitted [29, 2]. We use the idea of metaphor here rather broadly. A metaphor need not be precisely mapped from one discipline to another in order to be useful in that second discipline. For example, linguistic metaphors have been extremely powerful in suggesting lines of research

in unravelling the 'genetic code', even though they may not have been rigorously applied [24, 25 & 26]. Or again, as [22] and [48] show, Gregory Bateson drew heavily on metaphors derived from physics in developing the concept of 'feedback' which became central to the area of family therapy. As suggested in Figure 2, we do not know how important metaphorical understandings or adoptions are for knowledge processes or how common they are relative to other types of knowledge. Studies of knowledge transmission typically do not take into account this loose coupling of disciplines permitted by the exchange of metaphors. In our study, we will endeavor to locate such transmission through detailed content analysis of scientific papers, web sites, and personal interviews.

Sometimes metaphorical knowledge or information can be detrimental. For instance, Reddy [39] discusses the extension of the Shannon-Weaver (1948) information theory paradigm [46] to natural language communication. This extension suggests that language functions as a conduit, transferring thoughts bodily from speaker to listener. Reddy details various implications of the **conduit metaphor**, such as the assumption that thoughts and feelings can be ejected into an "idea space," become reified there, and later find their way into another mind. That linear account of communication does not account well for the active interpretation and reconstruction of meaning seen in many contexts of knowledge work.

Reddy questions the conduit metaphor: "I am suggesting, then, that... the conduit metaphor is leading us down a technological and social blind alley." (p. 310), proposing as one alternative, the **toolmakers paradigm**. The toolmakers paradigm itself is drawn from a metaphor in which one person creates a useful tool and attempts to communicate it to others. This communication is necessarily incomplete because the first person can never provide a total account of the relevant situational context for his or her tool. As a result, each additional person must reconstruct an image of the tool. Whether a conduit or toolmaker's metaphor is adopted has significant implications. For example, the conduit metaphor implies that successful communication with least effort is a desideratum, whereas the toolmakers paradigm implies that communication will always go astray unless real effort is expended to reach common understanding.

To understand various types of knowledge in an alliance, we view knowledge as highly contextualized, with the circumstances and conditions controlling how *knowledge claims* are adjudicated as further factors to be weighed as part of knowledge processes. Previous accounts of knowledge distribution and transfer begin with a thin conception of knowledge and the role of context in affecting the ways in which what stands as "knowledge" is understood and accepted as such. In our view, the contexts and practices in which knowledge claims are negotiated

must themselves be part of the process under study. These processes may be especially important in multidisciplinary teams of the type that we are studying. Hence, we have tried to go beyond unidirectional and simple models of knowledge distribution and transfer, to consider the rich and socially situated conditions of knowledge. We find this approach especially relevant to this study, because in exploring the ways in which alliances and other forms of collaboration use new technologies, we are also encountering, not *new* conceptions of knowledge (since many of these particular points have been made before by others), but new situated understandings of how knowledge *happens*.

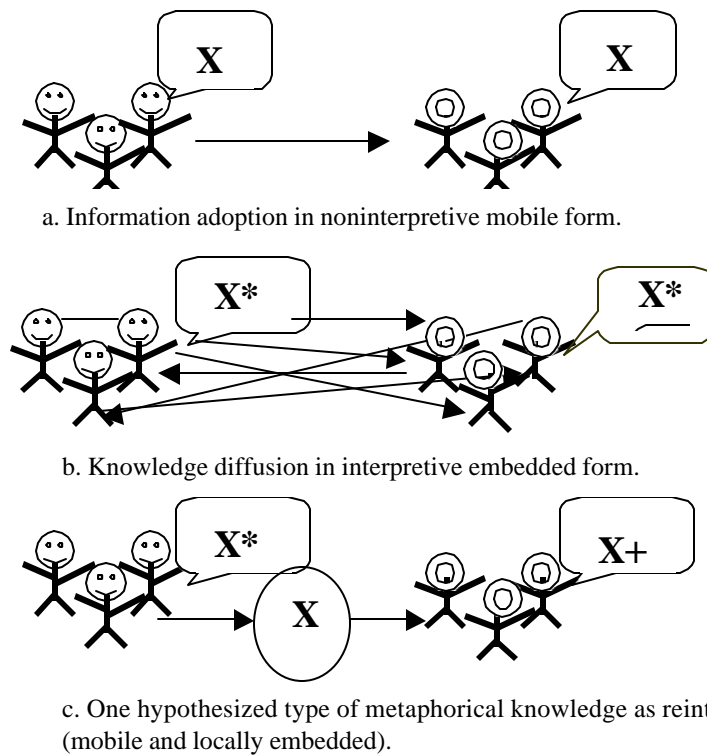


Figure 3. Potential models of interdisciplinary knowledge processes in alliances.

KNOWLEDGE IN ALLIANCES

Our overall project focuses on the role of a multidisciplinary alliance in the knowledge processes described above. Alliances often are unique in that they bring together participants from different disciplinary backgrounds, and the participants are often spatially distributed. Given these conditions, our research addresses how an alliance organization that is imposed on participants helps or hinders knowledge processes. How does an alliance provide

structure that may generate metaphorical thought? How does it provide a meeting ground or loci for scientific exchanges? How does one project serve as a metaphor or jumping off point for another?

Our discussion of metaphor bears resemblance to Star & Griesemer's (1989) concept of boundary objects [42], which are nomenclatures that make knowledge mobile, e.g., defining the names and contents of fields in a database. In multidisciplinary groups, we may see creation and use of boundary objects but what will be of interest to us is how group members came to the *definition* of the boundary object. Indeed, we are concerned as much with the variation in interpretation of such boundary objects or metaphors as with the similarities in how they are interpreted.

A schematic of proposed knowledge processes in multidisciplinary alliances is presented in Figure 3. We recognize at this point that the three configurations presented there do not represent all forms of knowledge exchange, but they represent three major configurations that we are using as a starting point for examining knowledge processes. In these representations, we go beyond previous work on knowledge by integrating into our conceptualization the role of electronic communication infrastructures and of group contexts in knowledge processes.

Each part of Figure 3 contains two teams of people each with different disciplines, group contexts or history. Lines of communication, computer-mediated *or* otherwise, between members of the two teams are illustrated. Our models hold that various knowledge processes are dependent upon these group contexts and communication patterns.

Imitative adoption: Figure 3a depicts how imitation or non-interpreted adoption can occur from one team to another based on minimal communication. Members of both teams identify the same term though without much complex comprehensive or thought to what it actually means, or how it might be related to a body of knowledge. This is similar to the conduit metaphor described above.

Supporting shared embedded knowledge: Figure 3b, on the other hand, suggests that for two teams that have very different contexts and history to come to common understanding of interpreted knowledge, there must be significant communication between team members to create a context that can support shared embedded knowledge. It may be that this type of sharing embedded knowledge across teams, and indeed even *within* multidisciplinary teams, as well as the mobile imitative knowledge are both relatively rare.

Sharing metaphorical representations of embedded knowledge: Figure 3c depicts the sharing of representations of knowledge embedded in either team (or, again, between members of a multidisciplinary team). Representations

make it possible for teams with different social, organizational and disciplinary contexts, and teams whose members are from significantly different disciplines to effect knowledge discovery by a synergistic pooling of their knowledge resources. It is by this type of exchange that we believe new knowledge is created in such alliances. This is similar to the toolmakers paradigm described above.

Overall, in the alliance organization, we expect to find the sharing of metaphorical representations of embedded knowledge as shown in Figure 3c to be the most common form. That is not to say that such knowledge processes that share metaphors across disciplines are not productive or authentic. The ability to skim a mobile metaphor off the top of a scientific literature may in fact spark more creativity within another discipline. In this manner, a multidisciplinary team member may actually contribute more new knowledge to his or her *own* discipline than to the multidisciplinary team. The contribution to the team is the individually or jointly constructed metaphor that is shared but put to different uses in different disciplines. This conceptualization of metaphorical knowledge sharing combines elements of embedded and mobile knowledge and may resolve some of the tension between the two.

These knowledge processes and various ways of sharing ideas across disciplines and across teams in an alliance may not be conscious. We expect to see various instantiations of the concept of “collaboration” with, in some cases, the adoption of an idea being unconscious and immediate, as if it were a self-evident truth. This is most likely for the metaphors that seem to have immediate meaning to all participants, albeit not always the identical meaning. In other cases we expect to see that the process is conscious, involving an unpacking of the idea to its parts, a deconstruction and reconstruction, a conscious attempt to interpret and evaluate the logic and applicability of the idea or theory. Where group members co-orient to a common intuitive metaphor, we may see immediate similarity in goals. Where such co-orientation takes a longer de- and re-constructive route we may see more discussion, less consensus, more negotiation and/or disagreement and argument.

Moreover, if the alliance as an organizational form matters, as opposed to these processes emerging naturally on their own, we would expect to see intense diffusion of ideas and the adoption and application of metaphors across teams both within the Alliance and back to the home disciplines of team members. We might also expect to see exchanges of information that are not immediately useful but that may expose individuals to new ways of thinking [15]. In addition, we would expect that alliances provide for bridging structural holes in knowledge structures [4], even if the players do not perceive a hole until a solution appears that bridges across disciplines. In the context of knowledge, metaphors are likely candidates for filling such holes.

RESEARCH SETTING: THE ALLIANCE

The National Computational Science Alliance (Alliance) is one of two partnerships in the National Science Foundation (NSF) sponsored Partnerships in Advanced Computational Infrastructure (PACI) which began in October, 1997. The Alliance has over 60 member institutions including, universities, national labs, companies, government agencies, community colleges and k-12 education partners. The mission of the Alliance is to "to prototype the computational and information infrastructure of the next century" (<http://www.ncsa.uiuc.edu/alliance/alliance>). Each year all partners compete for continuing funding, so organizations and individuals that are members of the Alliance will change over each of the ten years of the program.

An executive committee manages the Alliance with overall responsibilities for the distributed knowledge network and by several external advisory committees. Six Applications Technologies teams (AT teams) of scientists have been formed to use the infrastructures and identify further specifications for the advanced infrastructures. Three Enabling Technologies teams (ET teams) are architects of the infrastructures. In addition there are Education, Outreach and Training Teams, Partners for Advanced Computational Services and industrial and government partners and strategic vendors who participate in technology transfer.

The six AT teams in the Alliance (cosmology, environmental hydrology, molecular biology, chemical engineering, nanomaterials and scientific instrumentation) each are geographically distributed and the number of institutions ranges from five to nine per AT team. Currently the number of investigators funded on each of the AT teams ranges from 12 in environmental hydrology, to 6 in cosmology. While there are only a small number of funded investigators, each team has several unfunded collaborators, as well as a host of graduate students and post-docs working on the project. The teams were selected for participation in the PACI program based on their unique group and scientific histories, and on the degree to which they require various types of infrastructure. The goal was to include such a range of applications teams, that all computational, information and collaboration requirements would have to be prototyped. The teams do make differential use of communication technology. Some teams use the CAVE virtual reality environment, while others do not. The chemical engineering team reports their preferred mode of communication is telephone while the environmental hydrology team prefers email. Moreover, the AT teams are all multidisciplinary, some more than the others. For instance, while the team members have different expertise and skills, almost all scientists in the molecular biology team are in departments focussed on biology at the molecular

level such as biochemistry, molecular physiology, genome sequencing, with the additional expertise in computational biology and simulation cell biology. On the other hand, the environmental hydrology team is very interdisciplinary with scientists from atmospheric sciences, hydraulic research, computer science, systems ecology, environmental engineering, meteorology, space and science engineering, geography and computational fluid dynamics. The AT teams also vary in their group context. For instance the cosmology AT team grew out of a long standing grand challenge research group while the environmental hydrology group was formed for the Alliance by leading researchers who did not have a long collaborative history. While almost all of the molecular biology team knew each other before the Alliance, the network of the environmental hydrology (EH) team was not very dense at all, as presented in Figure 4 with pseudonyms.

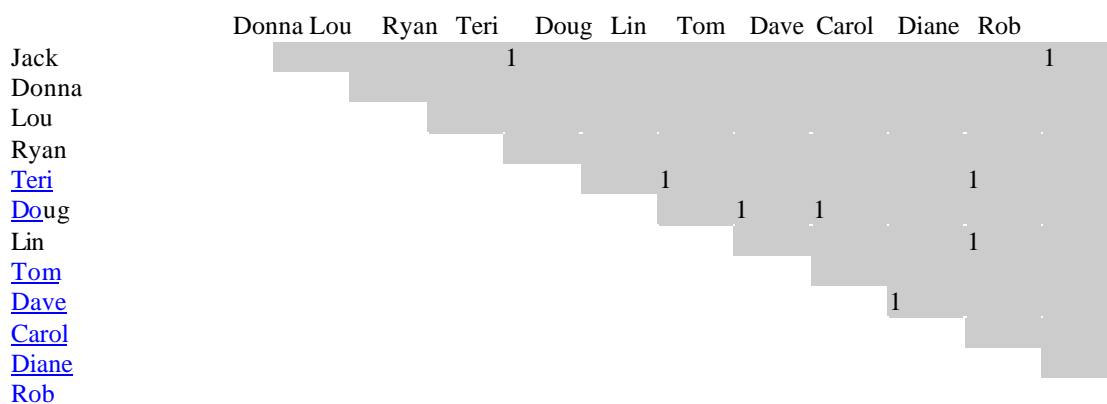


Figure 4. Who knew each other in the EH team before the Alliance
(Density = $8/66 = 0.12$)

An additional goal of the Alliance is to foster ties with others who are not funded by the Alliance to transfer knowledge gained from the Alliance to academia, government and industry. While the Alliance funds only a small number of AT scientists, each of them seem to have some strong ties outside of their team, naming anywhere from 3 to 11 collaborators outside of the Alliance. Additional influence of the Alliance can be seen in examining NCSA's client database, which includes the name, organization, status, project name and discipline for each person who has ever had an account on one of NCSA's computers. As of May, 1999 the client database contained 26,657 clients from 3,167 different organizations. In October, 1997 before the Alliance was formed, clients from only 1,136 organizations had accounts at NCSA. Nearly 12,000, or 44%, of all the clients were from organizations that are members of the Alliance (have someone funded by the Alliance at the organization). Moreover, the number of clients at an organization that was a member of the Alliance grew since the Alliance was formed. Based on initial

analyses using a fixed effects model it appears that joining the Alliance led to an increase in the number of clients from an organization. The average number of projects per client at an Alliance organization was significantly higher than the average number of projects for a client at a non-Alliance organization (2.96 vs. 1.63). The degree to which clients working on projects that are classified as a particular field of study are at Alliance organizations varies quite a bit as well. At the high end 89% of the clients working on a microelectronic information processing system project are at an Alliance organization and 88% from quantum electronics, waves and beams down to only 15% of the clients working in meteorology and 12% in computational mathematics

In addition to access to high performance computation to prototype multidisciplinary grand challenge applications, the AT scientists have access to an Alliance-wide intranet. Preliminary data collection indicates that the intranet is being used to support the Alliance as a networked organization. As of May, 1999 87% of the 507 intranet users were Alliance participants. Preliminary data also indicate that the different AT teams use the intranet as mode of communication to varying degrees. An estimated 40 intranet users worked on molecular biology projects while about half as many intranet users worked on chemical engineering projects.

So far, this section should have painted the skeleton of the group context, interdisciplinary nature, and use of the electronic infrastructure by AT teams. In addition, the scale of the Alliance and its impact has been outlined. Overall, there are three characteristics of the Alliance which are particularly important for our study: this alliance exemplifies many of the features of typical alliances, thereby providing an excellent case study for the potential strengths and weaknesses of the alliance model; this alliance has access to the latest technological resources to conduct leading-edge research, through its collaborative affiliations with institutions nationwide, which would allow the study of distributed knowledge infrastructures in their most sophisticated forms; this alliance has the explicit goal of developing and using advanced computational infrastructures to construct and disseminate knowledge in innovative ways; and finally, the members of this alliance also have ongoing collaborations with others outside of the Alliance. Hence the study of the Alliance and its infrastructures contributes to a “state of the art” understanding of these new knowledge infrastructures and their potential.

METAPHORS IN THE ALLIANCE

The multidisciplinary AT teams were brought together to solve problems, construct new knowledge, especially regarding computational science and disseminate it back to their home disciplines. In addition to the

within team, and within discipline knowledge processes, we have seen some evidence of metaphors crossing AT teams. A prime example is the "workbench" metaphor first conceived by an NCSA scientist in the Molecular Biology team in 1992. The idea behind the Biology Workbench™ was to include "all the software plus sequence databases available under a common Web Browser interface, so that the full functionality would be available on a platform-independent basis, even if the user had only a modest desktop machine. The first implementation of that, on an SGI machine set up as a Web server/analysis/search server in our Computational Biology Group, I believe was sometime in '94 or '95. The Workbench was made available to the public on an NCSA public machine in June '96" (Interview with Eric Jakobsson, May 1999).

The workbench metaphor caught on, and then-Alliance director, Larry Smarr began proposing NCSA Information Workbench as architecture for web based computing. NCSA publications claimed analogous workbench developments were underway. "In fact, the National Computational Science Alliance is already working on several toolkits that use the same interface as Biology Workbench, including the Chemical Engineering Workbench and the Environmental Hydrology Workbench" (ACCESS 2/9/99). Searching the NCSA Web site revealed 199 pages with "workbench." Yet none of those listings referred to an environmental hydrology workbench. And, when asked to describe how his team came up with the chemical engineering workbench, an NCSA chemical engineering AT scientist gave a long explanation, without any reference to the biology workbench. Clearly, there is some unconscious metaphoric learning and inventing in the Alliance. Thus the Alliance, with of six multidisciplinary teams and information and knowledge flowing within and between them via various media provides a rich research setting in which we can focus on the interaction between characteristics of knowledge, group context and infrastructure.

In terms of knowledge characteristics, our models connect with analyses of "mutual belief," from the analytical philosophy tradition, such as Stephen Schiffer's work on "meaning" [49, 50, see also 51]. Scholars in this area ask questions such as: Does it matter whether people are aware that they hold beliefs in common? Do jointly held beliefs function logically in the same way as individual beliefs? How do common beliefs permit or shape joint action? How do common beliefs arise? Analytical work in the area has strongly influenced research on multi-agent systems and psychological modeling. For example, work on "Interacting Plans" [52] examines shifting contexts of belief and how beliefs about the beliefs of others affect collaboration.

The range of technologies in use by the AT teams will allow us to conduct a study of how technologies are perceived and employed within the application teams. This study is based on taxonomy of technology applications that Bruce and Levin [53,54] developed <<http://www.ed.uiuc.edu/facstaff/chip/taxonomy/index.html>>. Its major categories are (1) inquiry, (2) communication, (3) construction, and (4) expression. Tools such as the radio telescope fall under a subcategory of #1, whereas typical CMC applications fall within #2. One goal is to see whether different teams can be productively described in terms of their differential use of and stance towards different tools and media. An hypothesis is that the function of key tools plays a major role in determining collaborative models. For example, the radio astronomy telescope serves a data-gathering function whereas the biology workbench serves a data-analysis function. These differences may help explain aspects of how team leaders interact with team members and with scientists in the larger community.

In addition, we will be looking for social network patterns that reflect the flow of metaphorical knowledge or interpretations of knowledge between different disciplines. We expect to find temporal stages in interaction beginning with first joining and learning to be comfortable with the infrastructure used to support the team, including learning the local social rules and behaviors. Then we expect efforts to be directed toward maintaining a presence in the multidisciplinary distributed research team. And, finally there might be a disconnecting; a withdrawal from the multidisciplinary distributed research team and rejoining or reconnected with the home discipline. Through these types of network movements, we will track the associated knowledge content shared. In the next section we describe the set of methods our team is employing to understand how knowledge processes can be distributed in a multidisciplinary alliance.

METHODS

The tension between the relative benefits and costs of embedded and mobile knowledge and the potential resolution of this tension with metaphors are especially significant in large-scale alliances. Such alliances are predicated on the assumption that a common information infrastructure will promote progress by enabling efficient communication and collaboration across distributed networks of people. However, the well-documented advantages of embedded practices localized in time and space suggest certain constraints on the efficacy of alliances as knowledge creating entities. These constraints and the compromises collaborators make in these sub-optimal distributed settings have not, however, been rigorously examined by social scientists.

Methods required to conduct such an investigation necessarily include a range of approaches involving various levels of participants in an alliance. The size of the Alliance with many participating organizations and hundreds of individuals, together with colleagues of AT scientists outside of the Alliance suggests that we need quantitative methods to assess infrastructure use and group context with knowledge construction. However, the complexity of an alliance structure demands that we also employ qualitative research methods to uncover the details of the interaction among these variables. Thus, both qualitative and quantitative methods are being used to map the communication channels within and between component organizations of the Alliance and with organizations outside of the Alliance, to assess their frequency of use, and to investigate how they shape the knowledge outcomes of an alliance, and are themselves influenced by various organizational, relational, and network context factors. Specifically, ethnography, situated evaluation, structured interviews, social network analysis, logfile analysis and document analysis methods each provide a unique approach to understanding the distributed knowledge processes in a multidisciplinary alliance.

Ethnographic studies of information systems provide in-depth understanding of work practices and the nature of information exchange between participants within such practices (e.g., Bowker, 1994; Star & Ruhleder, 1996; Bowker et al, 1997). The concept of boundary objects—objects which hold different meanings for different groups within an information system—has been highly influential in this field. Using boundary objects as a conceptual framework, ethnographic analysis can be used to look at the ways in which knowledge is both embodied - generated or captured in a local context within a person or given information system - and mobilized - moved from one context to another.

Situated evaluations typically focus on innovations as they are used across a variety of contexts (Bruce & Rubin, 1993). The key assumption is that any innovation comes into being only through its use. Using observations, interviews and surveys, situated evaluation examines the various realizations of an innovation in different settings. Its concern is with the characteristics of contexts that give rise to different realizations. This detailed articulation of the process whereby an innovation becomes realized in different ways can be useful in several ways, such as being able to study why the innovation was used the way it was in a setting.

Social network analysis provides a quantitative approach to analyzing social actors and patterns of relations among them, rather than focussing on individuals as a set of attributes (Freeman, Romney & White, 1992; Wasserman & Faust, 1994). Hence, social network analysis can be applied at every level of the study; for instance,

to investigate the patterns of relations among individuals, groups, and organizations. Social network data can be obtained through interviews and questionnaires. For instance, individuals can be asked questions such as: "Who do you work with?" "Who are your friends?," "Who do you go to for advice?," "Who do you go to for information about work-related techniques?" "How often do you communicate with each member of your group?" Network data can also be collected from organization charts, archival records, co-authorship, citation, co-citation and acknowledgements on publications (e.g., bibliometric methods), co-membership (e.g., in divisions, on projects, on committees), presence or participation in particular events (e.g., meetings).

Logfile analysis is a broad level view of who is using various technologies, how frequently and for what purposes. All server software applications, such as email, WWW and ftp, keep records of requests and actions taken in a logfile. The analysis of webserver logs, for instance, has become an important part of the Internet advertising industry (Riphagen & Kanfer, 1996), providing information about which files are requested most frequently, common paths through a website, where users link to a page from, as well as which domains users are coming from to access a website. If a website is password protected, the identity of actual users is recorded in the webserver log file.

Document analysis can be applied to traditional publications as well as to new styles of writing and sharing of texts via new technologies (Burbules & Bruce, 1995; Burbules, 1998). A range of quantitative and qualitative document analysis techniques can be employed at various levels of an alliance to understand the knowledge characteristics and group context. Qualitative document analysis will be used to investigate the details of knowledge such as when new ideas entered a group. Initial content analysis from a Toulmin perspective [55] suggests that Alliance scientists/technologists frequently employ explicit, anticipatory *warrants* in their writing as a means of persuading audiences--perhaps from other disciplines--to adopt their technological innovations. The notion of "warrants," refers to assumptions common to a discipline. In future studies, we will continue to use such discourse-based interviews to investigate whether introducing the team members to Toulminian analysis may enable them to communicate embedded knowledge more effectively to each other and to external audiences.

Bibliometric methods such as the analysis citations, co-citations, co-authorships, and co-word occurrences may be used on both hard copy and online publications (Callon, Law and Rip, 1986; Mullins, et al., 1977). Key measures include 1) total number of publications by team, 2) average number of authors per publication, 3) number of citations (i.e. times cited by others) per author and 4) total number of citations by team as well as 5) patterns of

collaboration evidenced in authorship and referencing, etc. Moreover, new electronic modes of producing and sharing documents provide data for event history analysis (Allison, 1984) of websites to investigate the rate of knowledge construction, and the amount of document sharing that occurs in distributed work groups. Initial analyses suggest that there are significant differences between teams and scientific disciplines in terms of multiple bibliometric measures such as coauthored publications, and co-citations (i.e. citing each others research). For example there is very little co-authorship among environmental hydrology team members while the cosmology team members have very high co-authorship rates as well as high individual publication rates. Moreover, the bibliometric analysis portrays cosmology as a relatively cohesive field in general.

Each of these methods are being brought to bear on the Alliance AT scientists, and their colleagues outside of the Alliance. In particular we are using these methods to attempt an understanding of the nature of the knowledge processes within the multidisciplinary teams, across the teams, and within each individual discipline.

DISCUSSION & SUMMARY

Currently we have just completed our first half-year of this three-year project. Our initial inquiry has been focussed on identifying our population of study as the exact nature of the AT teams involved is constantly evolving. Therefore we have conducted initial interviews with at least one member of each of the six AT teams. In addition we have constructed a large database of variables that help us to understand the relative degree of interdisciplinarity of the various teams. Moreover, as our research team has become distributed, rather than being located just on the UIUC campus, we have had to spend a bit of time testing and implementing videoconferencing equipment that facilitates our weekly meetings. As our project team grows, we find ourselves confronted with the same questions we ask of our research subjects about whether we can make mobile sophisticated methodological understandings that are embedded in our own various domains of inquiry.

In this very early stage of the investigation we are unpacking the concepts presented in Figure 1 which describes how group context, knowledge domain and communication infrastructure work together to influence knowledge production. The two-way nature of the relationships between these factors is emerging from our interviews with AT team members. Our original understanding of the communications infrastructure was how it was conceived to facilitate interactions and collaboration. However, the role of the technological infrastructure as a focus of the collaboration itself is emerging as a force that shapes the nature of the group.

Similarly, we have considered technologies as they enable knowledge discovery. But we are finding suggestions that the technologies also constrain the knowledge processes with computational models that contain implicit (or explicit) representations of the knowledge. For instance, initial ethnographic analyses of the Environmental Hydrology team at labs in Champaign, indicates that a key to understanding the general models being produced by the team is an understanding of the way in which the issue of boundaries (between river and bank, between air and earth) is articulated technically in the visualization and analysis programs being written.

Finally, we see various forms of collaboration or group context that are dependent on the nature of the disciplines, the types of problems the groups are solving, the state of knowledge in the field and the types of boundaries that have to be crossed in sharing knowledge and ideas. Relevant to Figure 3, we call into question the ideal notion of *collaboration*, and the normative force it holds. The type of group context may vary with respect to characteristics of who holds the knowledge. For instance it may be that in basic interaction, individuals own the knowledge and in collaboration groups hold the knowledge, while in a fully distributed team, no one person or group actually holds the knowledge. In this last situation, there is even a greater need to work together.

We believe that in a scientific age that seeks to solve highly integrative problems – such as that faced in Environmental Hydrology – and computational problems that necessarily involve collaboration between computer science and other knowledge domains, it is vital to understand how these teams work. We hope that the results of our work will help suggest where alliances may best be formed in the future, and to prepare future multidisciplinary teams for the types of collaboration and knowledge sharing processes they might expect.

REFERENCES

1. Badaracco, J. (1991). *The knowledge link: How firms compete through strategic alliances*. Cambridge, MA: Harvard Business School Press.
2. Bowker, G. C. 1993. 'How to be Universal: Some Cybernetic Strategies,' In *Social Studies of Science*, 23:107-127.
3. Bowker, G. C. & Star, S. L. (1999). *Sorting Things Out: Classification and its Consequences*, Cambridge, MA: MIT Press.
4. Burt, R. S. 1992 *Structural Holes, The Social structure of Competition*, Harvard University Press.
5. Carley, K., & Wendt, K. (1991). Electronic mail and scientific communication. *Knowledge*, 12, 4, 406-440.

6. Chubin, D. E., Porter, A. L., Rossini, F. A., & Connolly, T. (Eds.) (1986) *Interdisciplinary Analysis and Research*. Mt. Airy, MD: Lomond Publications, Inc.
7. Clarke, A. E. & Fujimura, J. H. (Eds.) (1992). *The Right Tools for the Right Job: At Work in 20th-Century Life Sciences*. Princeton, NJ: Princeton University Press.
8. Daft, R. L. & Lengel, R. (1986). Organizational information requirements, media richness, and structural design. *Management Science*, 32: 554-571.
9. Dede, C. (1996). The evolution of distance education: Emerging technologies and distributed learning. *American Journal of Distance Education*, 10(2), 4-36.
10. Duncker-Gassen, E. (1998). *Multidisciplinary Research at the University of Twente*. Enschede: Twente University Press.
11. Fujimura, J. H. (1998). *Crafting science: A sociohistory of the quest for the genetics of cancer*. Cambridge, MA: Harvard University Press.
12. Galegher, J. & Kraut, R. E. (1990). Technology for Intellectual Teamwork: Perspectives on Research and Design. In *Intellectual teamwork: Social and technological foundations of cooperative work*. Galegher, J., Kraut, R. E., Egidio, C. (Eds.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
13. Galegher, J., Kraut, R. E., Egidio, C. (Eds.). (1990). *Intellectual teamwork: Social and technological foundations of cooperative work*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
14. Galison, P. (1997). *Image and logic: A material culture of microphysics*. Chicago: University of Chicago Press.
15. Granovetter, M.S. (1973). The strength of weak ties. *American Journal of Sociology*, 78, 1360-1380.
16. Granovetter, M.S. (1982). The strength of weak ties: A network theory revisited. In P.V. Marsden & N. Lin (Eds.), *Social Structure and Network Analysis* (pp. 105-130). Beverly Hills: Sage.
17. Gulati, R. (1998). Alliances and Networks. *Strategic Management Journal*, 19, 293-317.
18. Hansen, M. T. (1999). The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*, 44, 82-111.
19. Haythornthwaite, C. (Jan., 1999). Collaborative work networks among distributed learners, *Proceedings of the 32nd Hawaii International Conference on System Sciences*. Los Alamitos, CA: IEEE Computer Society Press.

20. Haythornthwaite, C. (1996). Social network analysis: An approach and technique for the study of information exchange. *Library and Information Science Research*, 18, 323-342.
21. Haythornthwaite, C., & Wellman, B. (1998). Work, friendship, and media use for information exchange in a networked organization. *Journal of the American Society for Information Science*. In press.
22. Heims, Steve J.. The cybernetics group, Cambridge, Mass. : MIT Press, c1991
23. Hesse B.W., Sproull, L.S., Kiesler, S.P. and Walsh, J.P. (1993). 'Returns to science: computer networks in oceanography', *Communications of the ACM* 36 (8): 90-101.
24. Judson, H. F. (1996). The eighth day of creation : makers of the revolution in biology /, Horace Freeland Judson. Expanded ed. Plainview, N.Y. : CSHL Press.
25. Kay, L. E. (1993) The molecular vision of life : Caltech, the Rockefeller Foundation, and the rise of the new biology. New York : Oxford University Press.
26. Keller, E. F. Refiguring life : metaphors of twentieth-century biology /, Evelyn Fox Keller. New York : Columbia University Press, c1995
27. Kogut, B. (1996). What firms do? Coordination, identity, and learning. *Organizational Science*, 7(5), 502-519.
28. Koschmann, T. (Ed.)(1996a) *CSCL: Theory and Practice of an Emerging Paradigm*. Mahwah, NJ: Lawrence Erlbaum.
29. Lakoff, G. & Johnson, M. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
30. Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Cambridge, MA: Harvard University Press.
31. Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. New York, NY: Cambridge University.
32. Leonard-Barton, D. (1995). *The Wellsprings of Knowledge*. Cambridge, MA: Harvard Business School Press.
33. McGrath, J.E. (1984). Groups, interaction and performance. Englewood Cliffs, NJ: Prentice-Hall.
34. McPherson, J. M. & Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status, distance, and the composition of face-to-face groups. *American Sociology Review*, 52, 370-379.
35. Michaelson (a.k.a. Kanfer), A. G. (1993). The development of scientific speciality as diffusion through social relations: The case of role analysis. *Social Networks*, 15, 217-236.

36. Mullins, N. (1980). *Social Networks among Biological Scientists*. New York: Arno Press.
37. Pickering, A., ed. 1992. *Science as Practice and Culture*. Chicago : University of Chicago Press.
38. Price, D. J. (1986). *Little science, big science ... and beyond*. New York: Free Press of Glencoe.
39. Reddy, M. J. (1979). The conduit metaphor -- a case of frame conflict in our language about language. In Andrew Ortony (Ed.), *Metaphor and thought* (pp. 284-324). London, England: Cambridge University Press.
40. C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379-423 and 623-656, July and October, 1948.
41. Sproull, L. & Kiesler, S. 1991. *Connections: New Ways of Working in the Networked Organization*. Cambridge, MA: MIT Press.
42. Star, S. L., & Griesemer, J. (1989). Institutional Ecology, Translations, and Coherence: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-1939. *Social Studies of Science*, 19, 387-420.
43. Walsh J. & Bayma T. (1996) Computer networks and scientific work. *Social studies of Science*, 26, 661-703.
44. Walsh, J. Kucker, S, Maloney, N. & S. Gabbay, (1999). Connecting minds: CMC and scientific work. *Journal of the American Society for Information Science*. Available online at <http://www.uic.edu/~jwalsh/JASIS>
45. Wasserman, S., & Faust, K. (1994). *Social network analysis*. Cambridge: Cambridge University Press.
46. W. Weaver & C. E. Shannon, (1949). *The Mathematical Theory of Communication*, Urbana, Illinois: University of Illinois Press, republished in paperback 1963.
47. Wenger, E., (1998). *Communities of practice: learning, meaning, and identity*. Cambridge, U.K. ; New York, N.Y. : Cambridge University Press.
48. Wittezaele, J. (1992). *A la recherche de l'école de Palo Alto I*, Jean-Jacques Wittezaele, Teresa García. Paris : Editions du Seuil.
49. Schiffer, S. (1972). *Meaning*. Oxford: Oxford University Press.
50. Schiffer, S. (1987). *Remnants of meaning*. Cambridge: MIT Press
51. Meggle, G. "Common belief and common knowledge."
<<http://lucs.lu.se/spinning/categories/decision/Meggle/index.html>>.
52. Bruce, B. C., & Newman, D. (1978). Interacting plans. *Cognitive Science*, 2, 195-233
53. Bruce, B. C., & Levin, J. A. (1997). Educational technology: Media for inquiry, communication, construction, and expression. *Journal of Educational Computing Research*, 17 (1), 79-102.

54. Bruce, B. C., & Levin, J. A. (in press). Roles for new technologies in language arts: Inquiry, communication, construction, and expression. In J. Flood and D. Lapp, *Handbook of research on teaching the language arts*. New York: Macmillan.
55. Toulmin, S. (1958). *The uses of argument*. Cambridge, England: Cambridge University Press.
56. Inkpen, A.C. (1998). Learning and knowledge acquisition through international strategic alliances. *The Academy of Management Executive*, 12(4), 69-80.

ABOUT THE AUTHORS

Geoffrey C. Bowker is Professor in the Department of Communication at the University of California, San Diego. His most recent book is *Sorting Things Out: Classification and its Consequences* (Cambridge, MA: MIT Press, 1999).

Bertram C. Bruce is a Professor of Library & Information Science, and a Senior Research Scientist at the NCSA at the University of Illinois at Urbana-Champaign. His research focuses on information and communication technologies.

Nicholas C. Burbules is Professor of Educational Policy Studies at the University of Illinois at Urbana-Champaign and editor of the journal *Educational Theory*. His recent books include *Watch IT: The Promises and Risks of New Information Technologies for Education* (Westview), with Thomas A. Callister, Jr..

Caroline Haythornthwaite is an Assistant Professor in the Graduate School of Library and Information Science at the University of Illinois at Urbana-Champaign. Her research addresses problems of information exchange and community development in computer-mediated environments.

Alaina G. Kanfer is a consultant with BORN eBusiness Solutions Center and a Senior Research Scientist at the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign. She specializes in e-Commerce and business intelligence from a communities and exchanges perspective.

Joseph F. Porac is a Professor of Organization and Management in the Goizueta Business School at Emory University. His research interests include managerial cognition, the social construction of knowledge, and organizational impression management.

James B. Wade is an Associate Professor in the Management and Human Resources Department at the University of Wisconsin - Madison. His research interests include technology issues, corporate governance and organizational ecology.