A new type of scientific research group is emerging—the extended research group (a large number of geographically dispersed researchers working together in a coordinated fashion). This article describes the roles that various modes of communication (such as face-to-face discussion and EMail), types of information communicated, and types of social ties played in the communication behavior of one such group, computer analysts and programmers working on a specialized language model called Soar, which stands for "state operator and result." The authors show that the Soar group used different communication modes to acquire different types of information and that EMail enhanced rather than reduced the impact of traditional scientific communication networks.

Electronic Mail and Scientific Communication

A Study of the Soar Extended Research Group

KATHLEEN CARLEY
Carnegie Mellon University
with KIRA WENDT
Aion Corporation

Scientific progress depends on communication between scientists (Cole & Cole, 1973). As a consequence, the factors that affect communication, such as whether the mode is face-to-face or electronic, may influence research process and progress. Telecommunication technologies and especially electronic mail systems (EMail) may even change the nature of "doing science"

Authors' Note: This work was supported in part by National Science Foundation grant ASC-8617695 and by the Committee for Social Science Research on Computing at Carnegie Mellon University. Many thanks go to Allen Newell, John Laird, Paul Rosenbloom, and the members of the Soar extended research group for their cooperation, without which this study could not have been done. The authors would also like to extend their thanks to the anonymous reviewers for many insightful comments. A previous version of this article was presented at the annual meeting of the American Sociological Association, August 1988.

Knowledge: Creation, Diffusion, Utilization, Vol. 12 No. 4, June 1991 406-440

(Aborn, 1988). No longer, some observers claim, will researchers be engaged in science; rather, they will be engaged in telescience.[1]

Were the use of such communication technologies as EMail, facsimile transmission (FAX), and the telephone the only thing that characterized telescience, then telescience as a social phenomena would not be particularly interesting. The dominant difference between telescience and "conventional" science would be the speed of communication. Telescience, however, is characterized not only by the use of telecommunication technologies but by the proliferation of extended research groups who do science at a distance. As noted by Aborn and Thaler (1988),

> Increasingly, scientific explorations are being carried out by teams of researchers rather than single investigators, and more and more of those teams are composed of scientists from geographically dispersed sites. The ability to confer over long distances via any conceivable type of media, in real time as well as asynchronously, has thus created a considerable excitement among many in the scientific community. (p. 134)

In other words, science today is done not only by individuals but by research groups (a large number of researchers who work together in a cooperative and often collaborative fashion on a particular scientific problem under a single director, such as a faculty member and his or her core of graduate students). Such groups are housed within a single institution and may include researchers from multiple disciplines. Two examples are the Computer Aided Design Center at the University of California at Berkeley and the Computer Aided Design Center at Carnegie Mellon University. More recently, what appears to be a new type of research group is emerging—the extended research group, defined as a very large unified, cohesive, and highly cooperative research group that is geographically dispersed yet coordinated as though they were at one location and under the direction of a single director. Clearly, there are multiinstitutional coordinated groups that predate EMail, such as high-energy physicists working on accelerators. The extended research group appears to differ from these in its democratic division of labor, lack of competition for research funds, communication of new ideas prior to publication and before they are fully worked out, and unified perspective on what direction to move the research front. Our purpose here is not to determine whether this actually is a new type of group but to address the following question: *How do telecommunication technologies assist or make possible this new way of doing science (i.e., the extended research group)?*

There are only a few studies of the impact of telecommunication technologies or EMail on scientific communication (e.g., Freeman, 1984; Hesse, Sproull, Kiesler, & Walsh, 1990; Hiltz, 1978; Lievrouw, 1986, 1988; Rice &

Case, 1983). Furthermore, the communication activities of the extended research group have not been subjected to empirical investigation. In this article, we present an empirical study of the impact of EMail, relative to other modes of communication, on scientific communication among members of an extended research group. By EMail we are referring only to individual-to-individual or individual-to-group asynchronous computer-based communication and not to synchronous electronic communication systems, such as computerized conferencing. Admittedly, this study is highly exploratory and deals with only one extended research group, but we believe that our work has interesting implications for future work in the sociology of science and communication policy. The particular extended research group examined—the 26 dominant users of the computer architecture Soar—is of special interest because (a) it is one of the first such extended research groups, (b) Soar, has received national attention (Waldrop, 1988a, 1988b), and (c) Soar is currently being used in various countries around the world. Further, this group includes the three original developers (Allen Newell, John Laird, and Paul Rosenbloom), is spread across multiple institutions, and includes faculty, graduate students, and research staff.

To researchers in the artificial intelligence community, Soar is regarded, both as a general model of human cognition (Newell, 1990) and as a specialized computer language where each program is represented as a series of if-then rules (i.e., a production system) and the language has a built-in procedure for learning by creating new rules that encapsulate previous problem solutions (i.e., chunking; Laird, Newell, & Rosenbloom, 1987; Laird, Rosenbloom, & Newell, 1986a, 1986b). According to the Soar manual (Laird, 1986),

> Soar is an architecture for problem solving and learning, based on heuristic search and chunking. Soar is embedded in a production-system architecture—a modified version of *Ops5*—where all the volatile short-term information is held in *working memory* and all the fixed long-term knowledge is encoded as *productions*. (p. 1, emphases in original)
>
> Soar is an architecture for general intelligence that has been applied to a variety of tasks: many of the classic artificial intelligence (AI) toy tasks such as the Tower of Hanoi, and the Blocks World; tasks that appear to involve complex, non search reasoning, such as syllogisms, the three wise men puzzle, and sequence extrapolation; and large tasks requiring expert-level knowledge, such as the *R1* computer-configuration task. (p. 3, emphasis in original)

Soar is not a completed product; rather, it is still in the process of being developed and expanded. Nevertheless, as of January 1988, Soar was being used in 6 countries (United States, Sweden, England, Germany, Holland, and France), 14 U.S. universities, and at 20 industrial sites.

Our study focused on the period from 1986 to 1987 and used material from 1984 through 1988. As of June 1987, the Soar extended research group, the group of individuals who worked extensively with Soar and were involved in Soar's development, contained 26 members (18 graduate students, 5 faculty, 1 industrial affiliate, and 2 system programmers). These researchers were located at Carnegie Mellon (10), Stanford (9), Michigan (5), Digital Equipment Corporation (1), and in Europe (1). Most were members of or had degrees from computer science departments. At the time of our study, these individuals had been involved in the Soar project for times varying from a few months to 4 years.

The Soar extended research group thus represented a valuable opportunity for the study of scientific communication. First, because Soar was seen as either a theoretical construct (i.e., a model of human cognition or a unified theory of cognition) or as a technology (i.e., a production system language with learning), the group members communicated both technological information and theoretical information in the course of their work. Second, Soar's status at the time, as both under development and in use, required group members to communicate big innovative ideas as well as day-to-day information. Third, the Soar extended research group was composed of researchers at multiple institutions, distributed geographically, and with different substantive interests. This situation allowed us to examine the diffusion of information via different types of ties. Fourth, the extended research group was composed of graduate students and faculty who regularly used EMail and for whom EMail was not a novel toy. Hence our results may be applicable to other scientific fields in which EMail is becoming a mature technology. And finally, we were able to capture data on the diffusion of Soar as it happened and from almost the beginning of the diffusion process, rather than only after the fact (Rogers, 1983), providing insight into the process of scientific diffusion during the initial stages.

EMail and Scientific Communication

Previous research on the impact of EMail on individuals' communication behavior (for reviews, see Hiltz & Turoff, 1978; Kerr & Hiltz, 1982; Rice, 1980, 1984; Steinfeld, 1986) has shown that EMail eliminates the necessity of geographic or temporal synchronization (Hesse, Werner, & Altman, 1988; Huff, Sproull, & Kiesler, 1989), is fast and inexpensive (Hiltz & Turoff, 1978; Lievrouw, 1986), and affects communication efficiency by effecting group participation, interpersonal behavior, and group identification (Freeman, 1984; Kiesler, Siegel, & McGuire, 1984). Other work has shown that status differences are reduced when EMail is used (Sproull & Kiesler, 1986) and that EMail may affect innovation (Rice, 1987).

Within science, EMail has been expected for some time to affect the communication of information between and collaboration among scientists (Aborn, 1988; Hiltz, 1978; Hiltz & Turoff, 1978). For example, Rainwater and Smeeding (1988) noted that the ease of access to, and low cost of, telecommunication technologies are enabling "international collaborative research projects and colleagueship to reach their full potential" (p. 105). In another example, Bierly (1988) pointed out the importance of telecommunication technologies in studying climate and noted that the World Climate Program is a "prime example of gains in scientific knowledge achievable only through collaboration and communication on a worldwide basis" (p. 106). And finally, Hiltz and Turoff (1978) predicted that EMail (or what they initially called "computerized conferencing") would increase connections between members of invisible colleges and decrease the isolation of scientists who are in geographically remote institutions.

Research on scientific communication and the diffusion of information in general (Carley, 1990; Cole & Cole, 1973; Crane, 1970; Frydol & DeGreve, 1986; Knorr-Cetina, 1981; Menzel, 1960; Nelson & Pollock, 1970; Price, 1965, 1970) has suggested that individuals acquire information that is new to them (hence the information diffuses) from individuals with whom they share particular types of relationships or "ties." Two such ties identified as affecting the diffusion of information are invisible college ties (Crane, 1970; Price, 1963) and proximity ties (Barnlund & Harland, 1963; Festinger, 1950; Festinger et al., 1948). Although other ties and their patterns may be important (Allen, 1977; Menzel, 1960; Price, 1965), in our study, we examined whether EMail enhanced or reduced the effect of invisible college and proximity ties on the communication of scientific information.

Different types of information may also diffuse via different types of ties (Granovetter, 1973, 1974). In fact, a plethora of research illustrates that the diffusion of technologies and information about technologies (Anderson & Jay, 1985; Coleman, Katz, & Menzel, 1966; Johnson, 1986; Kara-Murza, 1981) and the diffusion of other types of information, such as gossip or theories (Allen, 1977; Cole & Cole, 1973; Festinger et al., 1948; Kara-Murza, 1981; Katz, Levin, & Hamilton, 1963) are affected by different factors. Whether these different factors extend to the technological and theoretical information that is being communicated via different ties is still an open question. We also do not know whether different communication modes are used to communicate technological and theoretical information. If different types of scientific information *are* communicated through different ties and via different communication modes, then, to the extent that EMail differentially affects these ties and replaces or is similar to those modes, EMail will differentially affect the communication of scientific information.

Plan of Study

By choosing three special factors – the mode of communication, the type of information being communicated, and the type of tie between people – we sought to determine whether EMail enhances the current communication system or has a reductionary or replacement effect.

Replacement and Enhancement Hypotheses

Some years ago, Thorngren (1977) noted that there were basically two lines of reasoning regarding the impact of telecommunications. According to his first line of reasoning, "Some see a new communications technology as almost automatically bound to produce new communication linkages between previously unconnected places and individuals." Nevertheless, he continued, "Others point out that without close 'fitness' to already existing structures, a new communication network will not long survive, and its development effects will be poor" (p. 374). At the time, Thorngren was talking about the impact of the telephone, but these same basic expectations exist for EMail. Since the first expectation is related to the expectation that old communication linkages will be reduced or even replaced by those formed through the new technology, we refer to this expectation as "the replacement hypothesis." The second expectation will be referred to as "the enhancement hypothesis."

Replacement Hypothesis

According to the replacement hypothesis, EMail, because of its characteristics, will replace traditional modes of communication such as face-to-face contact, telephone, and postal mail. Because EMail reduces physical and social communication barriers, it also should increase the opportunities for contact between strangers. These features of EMail, combined with the fact that EMail makes it easy to send mail to groups, lead to a variety of predictions for scientific communication:

1. Because EMail will reduce the use of, or replace, other modes of communication, all types of information are equally likely to be communicated electronically.

2. The ease of sending messages to both individuals and groups should lead to all information that is equally relevant to individuals and groups diffusing equally through individual and group EMail ties.
3. For a set of individuals, all of whom have access to EMail, the ease of communicating to groups and with strangers will lead to the creation of completely new communication groups (i.e., groups uncorrelated with groups that existed previously and used traditional modes of communication).
4. Given that individuals only have a limited time to communicate, the ease of communication with strangers and the existence of new communication groups will weaken existing ties between scientists. As a consequence of these changes, EMail usage should reduce the impact of proximity and invisible college ties on scientific communication. Historically, proximal and invisible college ties affected diffusion, and their absence enabled physical barriers to inhibit diffusion. Because EMail reduces physical barriers to communication, it should reduce the impact of both proximity and invisible college ties on diffusion.
5. Individuals who are tied, whether by proximity or an invisible college, often collaborate. EMail, by reducing the effect of proximity and invisible college ties, should shift the locus of collaboration. Hence collaboration should start occurring between pairs whose initial tie is through EMail and who are not, nor have been, proximal.

Enhancement Hypothesis

According to the enhancement hypothesis, EMail should complement traditional modes of communication if it is to survive. Most individuals who use EMail already know each other. Hence, even though EMail reduces social communication barriers and increases the opportunities for contact between strangers, such contact will be rare. These features lead to a variety of predictions for scientific communication:

1. Because EMail, relative to other communication modes, is not secure, because it increases the ease of group sending, and because there are few, if any, controls over resending messages, EMail may not be an appropriate mode of communication for certain types of information. Because EMail does not replace other modes of communication, the choice of mode depends on the needs and predilections of each researcher. For example, assuming that scientific reputations are most strongly built by creating new information, EMail's lack of security, its unofficial and lower status as a place of communication to other scientists makes it an unlikely choice for the diffusion of new primary information. Nevertheless, EMail and Group-EMail may be sufficient, or even preferable, communication modes for the diffusion of secondary information, bureaucratic messages, and repetitive detail.
2. Since EMail complements rather than replaces other modes of communication, diffusion may not occur electronically. Further, some types of information may be more likely to diffuse electronically than others.

3. Given the complimentary nature of EMail, the first communication groups to be established via EMail should be related to preexisting non-EMail communication groups.
4. EMail will strengthen or enhance existing ties because it can be used to maintain and nurture ties after they might have disintegrated, given only traditional modes of communication. EMail, like the telephone, reduces physical barriers to communication for individuals who were once proximal but are no longer and allows them to maintain prior levels of communication. It should be easier to maintain ties via EMail than the telephone because of EMail's lower cost and asynchronous nature. Therefore, because EMail reduces social, economic, and physical barriers to communication and encourages the creation of related communication groups, the use of EMail by members of an invisible college should strengthen those ties. Because Group-EMail can be sent to arbitrarily large groups, the size of invisible colleges can be increased beyond that sustainable through face-to-face communication. Thus invisible college and proximity ties, and not new EMail ties, should be the most common ties used for the diffusion of all scientific information.
5. Because EMail enhances the impact of invisible college and proximity ties, collaboration should be encouraged among individuals so connected. And, given EMail, collaboration that began when individuals were proximal should be easier to continue once the individuals are no longer proximal.

Data Collection

Data were collected on the diffusion and use of Soar within the extended research group during the period from November 1986 to November 1987. Most of the data are drawn from a questionnaire administered to group members at the third Soar workshop held in June 1987. All members of the Soar extended research group were present at this workshop and completed the questionnaire. This questionnaire included primarily subjective questions asking the individuals to recall how often they thought they used various modes of communication for communicating about what to whom, and when and how they first acquired information about Soar, its operationalization, and secondary information. Information on proximity and invisible college ties was also collected. During the months following, supplemental information was collected using a smaller questionnaire and by phone calls and EMail.

During this 1-year period, several things happened. The extended research group remained fairly stable – seven people entered and three people left, but these changes only involved graduate students. When Laird, a senior researcher, moved from Xerox Parc to Michigan, the extended research group spread from two to three schools. This change had little impact on personnel. Major changes in the theory underlying Soar and its architecture were discussed during this time period, although they were not implemented at this time. Diffusion of information about Soar and diffusion of the Soar program to outsiders steadily increased.

Mode of Communication

We collected data on what mode of communication the extended research group used to communicate what information to whom. The following communication modes were identified:

Face-to-face: One-on-one communication between proximal individuals
Telephone: Includes both one-on-one conversations and conference calls
Postal mail: Mail sent through U.S. postal service or other nonelectronic carrier
Other oral: Includes conferences, classroom interactions, group meetings, and so on
Other written: Includes articles, books, and magazines
EMail: Includes all electronic or computer mail regardless of source or destination.

EMail was further divided into three categories which differ in terms of the access and maintenance privileges of the reader: personal EMail, distribution lists (DLs), and electronic bulletin boards (b-boards). All forms are equally easy to use in terms of sending messages.

Personal EMail is EMail sent from one person to one other person, sometimes with copies to other persons. The receiver can choose to keep or delete personal EMail messages.

Distribution lists are a form of Group-EMail in which the sender sends a message to the DL and the message is then sent to every individual whose name is on the list. Each member of the DL gets his or her own copy of every message sent to the DL and can then choose to keep or delete it.

Electronic bulletin boards are a form of Group-EMail in which the sender send a message to the b-board and then anyone who has the correct access privileges can read the messages on the b-board. B-boards are much more open (i.e., more people can access the messages) than are DLs. Further, the individual reading the b-board can only read the message on the b-board and cannot choose to keep or delete messages. B-boards tend to be controlled by a b-board maintainer.

Types of Information

We categorized the information communicated by members of the Soar extended research group as follows. First, in terms of the work of doing science, there is a difference between primary and secondary information.

Primary information is the basic information that a new technique, model, or idea exists and is the type of information typically followed in diffusion studies. Such information often includes a very high-level or abstract description. Primary information is the knowledge claim. Secondary information is the detailed information that is needed for science to proceed and includes details on how to access a technology, the specific details on how the model operates, and coordination information. Coordination information includes information on which scientist does what and on when and where members of a scientific group will meet. Secondary information includes the important information needed to evaluate the knowledge claim. Thus secondary information is second only in the sense that its dissemination typically follows the dissemination of the primary information and not because it is less important. Second, in terms of information on Soar per se, there is a difference between the theoretical and the operational level of Soar.

We studied the diffusion of five different types of information related to Soar: two classified under primary information—big ideas and operational ideas—and three classified under secondary information—little ideas, operational details, and organizational information.

Big ideas. These are new theoretical models, approaches, paradigms, and so on. Examples of big ideas from three different disciplines are Soar as a unified theory of cognition (computer science), heat kills bacteria (biology), and electrons move toward a positive potential (physics).

Operational ideas. These are new technologies, software programs, mathematical models, and so on. Examples of operational ideas are Soar as a program, pasteurization, and tube diodes.

Little ideas. These include supporting theoretical arguments, confirming or disconfirming data, extensions of earlier approaches, paradigm details, and so on. Examples are automatic subgoaling, different bacteria have different shapes, and the greater the potential, the greater the speed with which electrons move toward it. Without these "little" ideas, the big ideas would be vacuous.

Operational details. These are the details on how the technology, program, or model works. Examples are specific algorithms, the specific temperature to which milk needs to be heated to kill bacteria, and the procedure for creating a vacuum inside tubes. Without operational details, scientists could not use a technology.

Organizational information. This is the information communicated between scientists for the express purpose of coordinating their research. Examples are meeting times, lists of group members, and delegation of responsibility.

Type of Tie

All members of the Soar extended research group clearly formed an invisible college, that is, they shared a common paradigm for doing research and had periodic face-to-face meetings (see Crane, 1970, for definition). All worked in artificial intelligence (AI) and used Soar to do research. They differed, however, in the problems to which they applied Soar (e.g., task acquisition, abstraction planning, syllogisms, or emotions). In terms of interaction, these individuals usually met about twice a year at the Soar workshops and at various AI meetings. Individuals at the same school met at weekly Soar meetings. Thus there are invisible college ties between all pairs of individuals.

Within the extended research group, however, some individuals share stronger invisible college ties than others in that they share more knowledge and have more face-to-face contact. We chose to examine mentor-student and co-student ties as the best indicators of these stronger invisible college ties. A mentor-student relationship exists between two individuals if one of the individuals listed the other as an academic adviser. A co-student relationship exists between two individuals if they were at the same school in the same department as students at the same time and had the same mentor. It is not necessary for the two students to be at the same point in the program (i.e., to be in the same cohort).[2]

The group members are also connected by proximity ties. To determine whether the individuals shared a proximity tie, each individual was asked to denote for every member of the user group whether they had an office next to each other, shared an office, lived in the same living group or apartment, were at the same school at the same time, or were in the same academic department at the same time. A proximity tie was placed between two individuals if any of these conditions were met.

Soar and the Extended Research Group

Soar was initially developed by Allen Newell in conjunction with two graduate students, John Laird and Paul Rosenbloom (Laird et al., 1986a, 1986b), who coined the term Soar, standing for "State Operator And Result," in winter 1982. The project began when Newell agreed to teach a course on artificial intelligence. To prepare for the course, Newell wanted to build a system that allowed users to write directly in problem spaces, and so he asked John Laird to develop a task experimentation system in the computer language LISP. Although this system never worked, the two researchers tried the same basic idea again in 1983 using production systems. This idea was initially called SOR and, later, Soar. In the summer and fall of 1983, the interests of Newell, Laird, and Rosenbloom merged—Rosenbloom started working on the reimplementation of R1[3] in Soar and Laird started working on the implementation of learning in Soar. Table 1 shows a time line for important Soar events.

Laird, Rosenbloom, and Newell started working extensively on the Soar project in the summer of 1983. At the time, Newell was having back problems, so the group met at his house for several hours every week. In the fall, Laird and Rosenbloom's offices were next door to each other, so they communicated by "popping in" to each other's offices. They had weekly 2-hour meetings with Newell. They also used EMail extensively. In 1983, Soar was first formally presented to scientists outside the development group. In the summer of 1984, after graduation, Rosenbloom went to Stanford and Laird to Xerox Parc. EMail usage increased. Newell flew to California about every 3 months. By January 1988, the three developers were holding regular weekly 2-hour telephone conference calls.

By spring 1986, the number of researchers working intensively with Soar and abetting its development had grown and the developers believed it necessary to hold a meeting at which current research on Soar could be presented and changes to the architecture discussed. Attendance at the Soar workshop was limited to the extended research group and a few special invitees.

The strong invisible college ties, for the most part, appeared to form a hierarchy (see Figure 1) because Soar was so new that it was still largely being worked on only by Newell and his students and their students. Only 5 of the 26 individual users did not share mentor-student or co-student ties to any of the individuals in the group (see box in Figure 1).

The extended research group was also fairly densely connected by proximity ties (see Figure 2), again a product of the newness of Soar, and the fact that most of the extended research group were at three universities. There were, however, many cross-university proximity ties because the principle actors moved and graduate students often spent summers or school terms at other universities.

We cannot completely disaggregate the impact of the invisible college and of proximity because we are convinced we were witnessing the birth of an

TABLE 1
Soar Time Line

Data	*Event*[a]
1982 Fall-1983 Spring	Soar developed
1983 Summer	First general public presentation of Soar at the International Joint Conference of Artificial Intelligence (IJCAI) First conference publication
1983 Fall	Carnegie Mellon Soar meetings begin
1984 Spring	Soar manual becomes available
1984 Summer	Second general public presentation of Soar at the American Association of Artificial Intelligence (AAAI) Laird moves to Xerox Parc Rosenbloom moves to Stanford University
1984 Fall	Stanford Soar meetings begin
1985 Summer	First refereed publication – *Institute of Electronic and Electrical Engineers: Pattern Analysis and Machine Intelligence* Newell gives talk at Cognitive Science Workshop, Irvine, CA Laird gives talk at Machine Learning Workshop, Skytop, PA
1986 Spring	Soar Workshop 1 is held *Machine Learning* journal article published Soar manual becomes a technical report
1986 Summer	First major book on Soar appears Chunking journal article appears in *Machine Learning*
1986 Fall	Soar Workshop 2 is held Carnegie Mellon hackers meetings begin
1987 Spring	Unified Theories of Cognition lectures are given at Harvard University *Artificial Intelligence* journal article published
1987 Summer	Soar Workshop 3 is held Laird moves to Michigan
1987 Fall	Unified Theories of Cognition lectures are given at Carnegie Mellon University Carnegie Mellon Unified Theories of Cognition meetings begin Rosenbloom moves to Los Angeles
1988 Spring	Soar Workshop 4 is held

a. Only major events in the early diffusion and development of Soar are listed.

invisible college. As a consequence, there was a high overlap between invisible college and proximity ties. Thus, in this article, we will not be concerned with whether invisible college or proximity ties have more of an

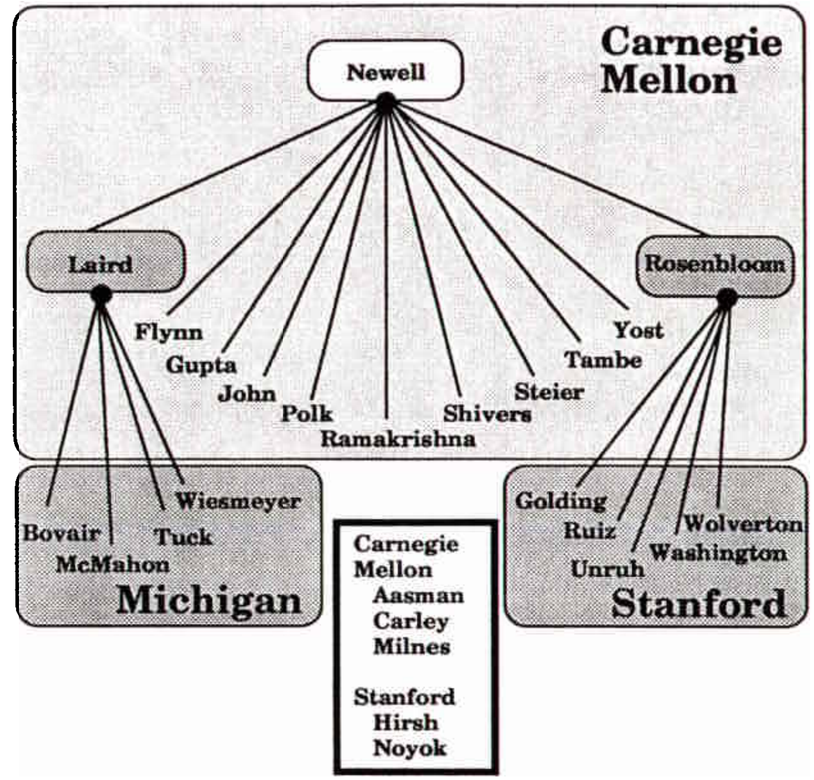

Figure 1: Soar's Invisible College
NOTE: This is the network formed using the strong invisible college ties. A hierarchy forms among most researchers because the invisible college is so new that it is still primarily formed of Newell's students and their students. A line indicates that there is a student-mentor or student-student relationship. In addition, individuals are marked by their current or most recent university affiliation.

effect on acquiring scientific information but on whether EMail affects the use of such ties.

EMail and the Soar Extended Research Group

The 26 members of the Soar extended research group were well acquainted with EMail; all but 2 had EMail IDs and even they had previous experience with EMail. There was one Group-EMail facility for the entire

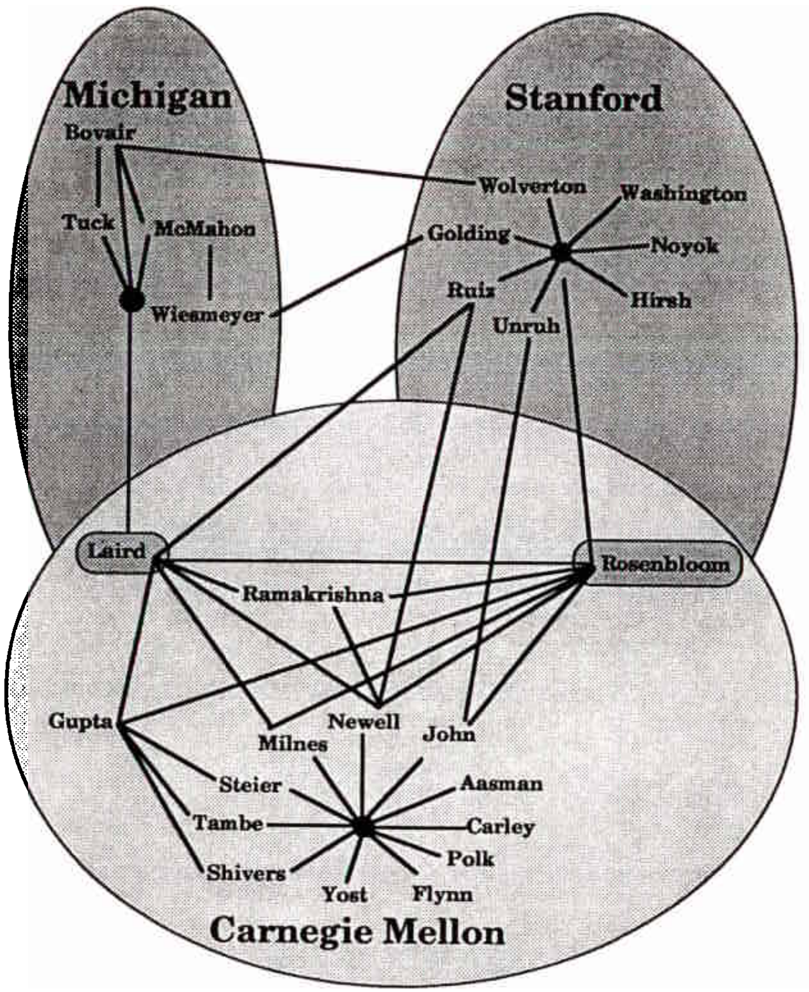

Figure 2: Proximity of the Dominant User Group
NOTE: This is the network formed using proximity ties. The entire extended research group is densely connected through proximity. A line indicates that the two connected individuals are proximal.

Soar extended research group, the Soar distribution list, and all members of the extended research group were on this DL.

We asked a sample of these individuals to track their EMail usage during a typical week.[4] Based on that data, we concluded that not only was EMail

TABLE 2
EMail Usage

~60	EMail messages sent and received (per week)
~12	Soar personal EMail messages sent (per week)
~24	Soar personal EMail messages received (per week)
12	Soar distributor list: Number of messages (per week)

TABLE 3
Breakdown of Soar Distributor List Messages

Year	*Primary Information*	*Little Ideas*	*Operational Details*	*Organizational*	*Other*	*Total Messages*
1984	0	0	0	5	1	5
	0%	0%	0%	100%	20%	
1985	0	9	22	54	13	86
	0%	10%	26%	63%	15%	
1986	0	26	55	130	41	215
	0%	12%	26%	60%	19%	
1987	2	15	75	257	66	376
	1%	4%	20%	69%	18%	

NOTE: Cell values are the number of messages containing that type of information. The percentages are the number of messages of that type as a percentage of the total number of messages during that year. A message containing more than one type of information is coded as a separate message for each type of information.

used frequently but that over half of the group's EMail had to do with Soar (see Table 2).

The members of the Soar group used personal EMail and Group-EMail to communicate several types of information. Table 3 breaks down all the Group-EMail by year, from August 1984 through December 1987.[5] The dominant use of Group-EMail was to communicate secondary information (i.e., to communicate little ideas, operational details, and organizational information). The second main use of Group-EMail was to communicate social information (classified as "other"). It was extremely rare for Group-EMail to be used for the communication of primary information.

Over 60% of the Group-EMail messages was organizational. An example of such a message is the following:

> Copies of "SOAR: An architecture for general intelligence" by Laird, Newell & Rosenbloom are available in 5216 WeH.

Over 20% of the Group-EMail messages dealt with operational details. The following sample message contains both operational details and organizational information:

> Cmu Soarers,
>
> I have run up against a bug, which I believe to be in cmu rt lisp, which causes the image that I released before to fail after a call to restart-soar.
>
> So in case any of you use restart-soar, or do not load the default productions as they stand, I am now releasing two separate images.
>
> They are on GEO and the H in /usr/milnes/soar/soar4-5/src/ and are named soar, for just the soar code and no default productions, and soar+default for soar with the default productions.
>
> If you would like the latest style default productions, and I have not changed them so they are still as on /h/usr/gca/v44/csoar/default.soar, they are also now in /h or geo /usr/milnes/soar/soar4-5/prg/default/new/default.lisp.
>
> Brian
>
> P.S. As I am now convinced that we need a soar-cmu mailing list I'm creating one now.

And finally, just under 20% of the messages were not direc.ly work related but were social, for example, containing information on transportation arrangements.

These data suggest that in a community where EMail is not a novel phenomena, Group-EMail may be used by scientists and researchers primarily to maintain the invisible college. For a scientist to remain an active member of an invisible college, the scientist needs access to organizational information. Further, when the members of the invisible college are using a common technology, individuals are more likely to maintain their membership in the group if they have access to the technology and it is easy to use. Thus Group-EMail serves an important role in science in efficiently providing all invisible college members and a large number of others with vital organizational and operational information. EMail may be replacing the traditional interoffice memo within the organization. But it also appears to be extending the functionality of this memo between organizations by allowing researchers at several geographically dispersed organizations to coordinate as if they were in one organization and in the same location.

The foregoing analysis focused on the number of messages and not the length or perceived importance of messages. Based on a sample of messages, however, it appears that messages containing "little ideas" are, on average, the longest messages sent, whereas organizational messages are the shortest.[6] This finding suggests that while the primary use of Group-EMail, in terms of number of messages, may be to maintain invisible college ties, the most

TABLE 4
Communication Modes Used to Communicate Information on Soar

Communication Mode	*Request*	*Give*	*Both*
Face-to-face	20 76.92%	18 69.23%	17 65.38%
Conference	14 53.85%	13 50.00%	12 46.15%
Telephone	11 42.31%	10 38.46%	8 30.77%
Postal mail	3 11.54%	6 23.08%	2 7.69%
Personal EMail	19 73.08%	14 53.85%	13 50.00%
Soar Group-EMail	9 34.62%	10 38.46%	6 23.08%
Electronic B-Boards	0 0.00%	0 0.00%	0 0.00%
N	26	26	26

important use of Group-EMail, in terms of message length, may be to nurture the invisible college and to increase its cognitive homogeneity by providing stimulating new information to all group members.

EMail was not the only mode of communication used by members of the Soar group. Group members also used face-to-face communication, telephone, group meetings, and even postal mail. We asked the members of the extended research group to check all communication modes that they used to (a) request information on Soar or (b) give information on Soar (Table 4). We found that more individuals used one-on-one, face-to-face communication than any other mode to both give and request information on Soar, but then tended to choose personal EMail, conferences, telephone, and Group-EMail in that order. The difference between face-to-face communication and EMail was not significant. A significant difference between EMail and either postal or telephone usage, however, suggests that although EMail enhanced rather than replaced face-to-face exchanges, it may have replaced telephone and postal exchanges.[7] Only a few individuals used postal mail. Further, there was a slight difference between synchronous and asynchronous modes of communication, with synchronous modes of communication (face-to-face, conference, and telephone) showing less discrepancy between who requests and who gives information than there was for the asynchronous modes of communication. In particular, whether because personal EMail reduces status queues or because it is easy to use, these individuals seemed more likely to request than to give information by EMail. Assuming that these requests are answered, this result suggests that EMail will facilitate greater understanding and collaboration by making individuals feel more free to gather information. Further, EMail may be replacing the telephone and postal mail only as a mode for requesting information.

No one used general electronic b-boards to communicate Soar information, that is, those b-boards not associated with the Soar project, as the developers had deliberately refrained from putting information about Soar on the general electronic b-boards lest it be prematurely viewed as a product available to the public. All other members of the Soar extended research group followed the lead of the developers in this matter. As Laird noted during our winter 1987 interview with him:

> I should probably tell you that we don't really want a fast dissemination to the masses at this point, and that is why we have not posted on b-boards or participated in computer conferences. At the talks, we do not say that Soar is available on request, since we are not equipped to handle 50 requests at once. We are really trying to control the diffusion—you tell people you know, but no advertising to the masses.

Information about Soar was diffusing outside the group, however. Presentations related to Soar at general artificial intelligence conferences ensured diffusion of at least the theoretical material. Even the Soar code was given to outsiders, although dissemination of that code was still highly controlled.[8]

The members of the Soar group also were members of multiple communication groups. By January 1988, there were 12 special formalized communication groups (see Table 5). Each group was characterized by having a dominant mode of communication and a dominant topic of discussion.[9] One communication group was composed of the three principal developers—Newell, Rosenbloom, and Laird—who held weekly phone conferences. Two other groups at Carnegie Mellon University and Stanford University had weekly meetings, usually run like small seminars at which members presented their research. Associated with the Carnegie Mellon and Stanford groups was a DL group, where the DL was used primarily to communicate information on meeting times, location, and topic. At the time of our survey, Laird had just moved to Michigan, so the Michigan group was in its infancy; yet by June 1988, there were weekly Michigan group meetings and an associated group DL.

TABLE 5
Communication Groups

Name of Group	*Communication Mode*	*Main Topic*
Developers	Telephone	Soar development
Carnegie Mellon Soar group	Oral group meetings	Soar and applications
Carnegie Mellon Soar DL	Group-EMail	Organization of group
Stanford Soar group	Oral group meetings	Soar and applications
Stanford Soar DL	Group-EMail	Organization of group
Soar extended research group	Workshops	Soar and applications
Soar extended research group DL	Group-EMail	Soar information
Active Soar users	Group-EMail	Soar bugs
Unified Theory of Cognition	Oral group meetings	Soar—unified cognitive theory
Unified Theory of Cognition DL	Group-EMail	Organization of group
Parallel Processing	Oral group meetings	Implementing parallel Soar
Parallel Processing DL	Group-EMail	Organization of group

Even by January 1988, all members of the Soar group (who had access to EMail) were in a single communication group connected by the general Soar DL. All active users reported bugs in Soar by using personal EMail to send Soar bug reports to a "fake person"—Soar-Bugs. These bug reports were then perused by the programmers, discussed with the developers, and, if necessary, put in a queue to be fixed. In addition, at Carnegie Mellon, there were the Unified Theories of Cognition and Parallel Processing subgroups. These groups held irregularly scheduled meetings that were run like small seminars. Associated with each of these two subgroups was a DL group. And finally, all members of the Soar group were connected by Soar workshops. By June 1988, four such workshops had been held, one each in spring and fall 1986, summer 1987, and spring 1988. These workshops were run like miniature conferences, with only one session at a time and with each attendee presenting information on his or her current research.

In January 1988, half of the communication groups used EMail as the dominant communication mode. Five of the non-EMail groups had a corresponding EMail group. This finding reinforces the earlier point made that one of the primary, and important, uses of Group-EMail in science is to maintain and enhance the impact of other communication ties. In the case of the Soar group, EMail was not used to create new ties but to maintain existing invisible college and proximity ties.

Diffusion and the Dominant Users

New Soar-related information diffused to members of the Soar group via many different communication modes. Further, which mode was used depended on what information was being communicated and how the individuals were tied.

Diffusion of Primary Information

Members of the Soar group were asked to denote, for primary information, both the type of contact from whom they learned that information and the communication mode used. To establish a context in which the data on the diffusion of primary Soar information could be interpreted, we gathered data on the diffusion of primary information on two items intimately related to Soar: "chunking" and "problem solving as search through problem spaces (search)." Both items have a theoretical and an operational component.

Both research areas had been developed more than 10 years before Soar, yet their implementation in Soar had been somewhat unique. Hence, if the length of time that information has been available influences its mode of communication, or the type of contact used for diffusion, then the diffusion pattern for big ideas should be different than for operational ideas.

Members of the Soar group were asked how they first heard the terms (and hence theoretical information about) for three concepts—"Soar," "chunking," and "problem solving as search through problem spaces"—and were also asked by which communication mode they first heard of the Soar-based operationalization for each of the same three items (see Table 6).[10]

In the survey instrument, members of the extended research group could choose among the following communication modes: face-to-face, telephone, conference, postal mail, personal EMail, Soar Group-EMail, electronic b-boards, and "other." No one, however, reported initially learning either the big ideas or the operational ideas by Soar Group-EMail, electronic b-boards, or postal mail. If the respondent listed "conference" or under "other" listed a verbal communication mode, then this response was classified as "other oral." When the respondent listed a written communication mode under "other," it was tabulated as "other written." One member of the extended research group could not remember by what mode of communication he learned the term "problem solving as search through problem spaces."

We conclude that individual scientists may be most likely to hear about big ideas first through one-on-one, face-to-face communication if the idea is a recently developed idea. Within science, once the idea has been established, however, one-on-one, face-to-face communication may be no more commonly used for diffusion than are other oral modes of communication, including one-on-many situations such as classrooms and conferences (the differences between face-to-face and other oral for chunking and search were not significant).[11] Even established ideas seem unlikely to diffuse initially

TABLE 6
Diffusion of Big Ideas and Operational Ideas, by Communication Mode

Communication Mode	*Soar*	*Chunking*	*Search Through Problem Spaces*
Big ideas			
Face-to-face	19	10	11
	73.08%	38.46%	44.00%
Telephone	1	0	0
	3.85%	0.00%	0.00%
EMail	0	0	0
	0.00%	0.00%	0.00%
Other oral	6	12	12
	23.98%	46.15%	48.00%
Other written	0	4	2
	0.00%	15.38%	8.00%
N	26	26	25
Operational ideas			
Face-to-face	13	11	10
	50.00%	42.31%	38.46%
Telephone	0	0	0
	0.00%	0.00%	0.00%
EMail	0	1	0
	0.00%	3.85%	0.00%
Other oral	4	5	9
	15.38%	19.23%	34.61%
Other written	9	9	7
	34.61%	34.61%	26.92%
N	26	26	26

through written modes of communication. One must be careful in generalizing this finding, however, because most scientists in our study were "new scientists" (i.e., graduate students or recent Ph.D.s). Whether reliance on oral rather than written communication extends to more established scientists is a point for future research.

For operational ideas, most diffusion appears to have occurred through oral modes of communication, especially one-on-one or face-to-face. Written modes, however, remain important as a mode for communicating operational ideas. Once individuals find out about the big idea, then they may be more alert to operational ideas contained in secondary information.

These data suggest that EMail was not used for the diffusion of primary information. No member of the Soar extended research group learned about any of the big ideas through EMail, either personal or group. And only one individual learned about an operational idea, "chunking," electronically.

A comparison of the data shown in Table 6 reveals a timing effect. The older theoretical ideas—"chunking" and "problem solving as search through problem spaces"—appear to have exhibited a different diffusion pattern than did the newer idea—"Soar." At the operational level, all ideas appeared to have been similar in age, and all operational ideas diffused via a similar pattern.

Members of the Soar group were asked to denote the type of contact from which they first heard the term for each of the three items—"Soar," "chunking," and "problem solving as search through problem spaces"—and then denote the type of contact from which they first heard of the Soar-based operationalization for each of the same three items. Responses to both questions are shown in Table 7.

The most common diffusion contact for big ideas appears to have been a student's professor. For both "Soar" and "problem solving as search through problem spaces," individuals tended to learn the concept first from a faculty member, typically during a one-on-one, face-to-face meeting. For the more established items, the initial contact was also a professor, but in a classroom setting. Indirect contacts (e.g., reading an article) accounted for little diffusion.

For operational ideas, the most common contact was indirect; the next most common contact was a professor in a one-on-one meeting. When we questioned members of the Soar group further, we found that invisible college and proximity ties were used to acquire the article. The receivers did not just stumble over the article containing the operational idea but were given it or told to read it by someone with whom they shared either an invisible college or proximity tie or both.

TABLE 7
Diffusion of Big Ideas and Operational Ideas, by Type of Contact

Type of Contact	*Soar*	*Chunking*	*Search Through Problem Spaces*
Big ideas			
Friend	3	5	0
	11.54%	19.23%	0.00%
Fellow worker	2	0	0
	7.69%	0.00%	0.00%
Student	1	0	0
	3.85%	0.00%	0.00%
Professor	12	5	11
	46.15%	19.23%	44.00%
Class	4	10	10
	15.38%	38.46%	40.00%
Conference	0	1	1
	0.00%	3.85%	4.00%
Article	0	4	2
	0.00%	15.38%	8.00%
Other	4	1	1
	15.38%	3.85%	4.00%
N	26	26	25
Operational ideas			
Friend	2	2	1
	7.69%	7.69%	3.85%
Fellow worker	1	2	1
	3.85%	7.69%	3.85%
Student	1	0	0
	3.85%	0.00%	0.00%
Professor	8	7	8
	30.77%	26.92%	30.77%
Class	2	2	4
	7.69%	7.69%	15.38%
Conference	0	1	2
	0.00%	3.85%	7.69%
Article	9	9	7
	34.62%	34.62%	26.92%
Other	3	3	3
	11.54%	11.54%	11.54%
N	26	26	26

TABLE 8
Diffusion of Big Ideas and Operational Ideas by Tie Used

Tie	*Soar*	*Chunking*	*Search Through Problem Spaces*
Big ideas			
Invisible college	14 53.85%	5 19.23%	21 84.00%
Proximity	10 38.46%	15 57.69%	10 40.00%
Combined	22 84.62%	20 76.92%	22 88.00%
N	26	26	25
Operational ideas			
Invisible college	9 34.62%	9 34.62%	9 34.62%
Proximity	6 23.08%	6 23.08%	6 23.08%
Combined	14 53.85%	13 50.00%	14 53.85%
N	26	26	25

Although the data on type of contact do show some timing effect, it was not as consistent as for mode of communication. There was no difference for operationalized information; for theoretical ideas, Soar was dominated by the professor contact.

The members of the Soar group were also asked to note the name of the person from whom they first heard the term or learned the operational idea. Table 8 compares diffusion ties with invisible college and proximity ties. Group members could learn these concepts from people outside of the group as well as through invisible college and proximity ties, thus reinforcing the impact of these ties beyond that shown here. Table 8 indicates that primary information, especially big ideas, may diffuse through either invisible college or proximity ties, regardless of the age of the item diffusing.

Diffusion of Secondary Information

Secondary information, unlike primary information, is composed of many more pieces of information. Thus we did not look at the tie or mode of communication used to learn a specific piece of information as we did in the case of primary information. Rather, we looked at the type of tie or mode of

communication used in general to acquire secondary information. For example, the data in Table 3 are predominantly secondary information.

Internal Communication Structure

Members of the Soar group were asked to denote what percentage of their interaction with each other member was by which mode of communication. We then constructed a set of five matrices for face-to-face, telephone, postal mail, EMail, and Group-EMail. A cell entry (Individual i × Individual j) is the frequency with which Individual i indicates that he or she interacted with Individual j using that communication medium (Figure 3). Note that it is possible for these matrices to be asymmetric. To extract a communication structure for this group, we ran the CONCOR program (Breiger, Boorman, & Arabie, 1975) on all five matricies at once.

The groups labeled Carnegie Mellon and Stanford in Figure 3 contain primarily people who were at each university in summer 1987. The third group (developers) contains developers and programmers who maintained and developed Soar as a technology. The fourth group (theoretical) contains Newell and other people involved in the development of Soar but who worked at a theoretical level. And the final group (no EMail) contains the only members of the extended research group, both at Michigan, who did not have access to EMail at the time of the study.

In Figure 3, we can see that all members claimed to communicate using a face-to-face mode, except for the two Michigan students who did not have access to EMail. They claimed not to interact face-to-face with people at Carnegie Mellon and Stanford, despite having attended a Soar workshop. Except for occasional phone calls, the only groups that regularly used the telephone to communicate were the members of the two development groups (primarily the weekly phone conferences between the three original developers). Only the individuals who did not have access to EMail used postal mail. The two university groups used EMail mainly to communicate with the developers, rather than with each other or among themselves. In contrast, the developers, especially the theoretical development group, used EMail to communicate with everyone who had access to EMail.

Group-EMail use showed opposite patterns. The developers used Group-EMail primarily to communicate internally, while the Carnegie Mellon and Stanford groups used Group-EMail to communicate with the development group. Only the Carnegie Mellon group used Group-EMail to communicate with the theoretical development group (which included Newell). These seemingly odd Group-EMail findings were the result of the individuals' *perceptions* of whom they communicated with using Group-EMail, not a

Face-to-Face

	C	S	D	T	N
CARNEGIE MELLON (C)	1	1	1	1	1
STANFORD (S)	1	1	1	1	1
DEVELOPERS (D)	1	1	1	1	1
THEORETICAL (T)	1	1	1	1	1
NO EMAIL (N)	0	0	1	1	1

Telephone

	C	S	D	T	N
CARNEGIE MELLON (C)	0	0	0	0	0
STANFORD (S)	0	0	0	0	0
DEVELOPERS (D)	0	0	1	1	0
THEORETICAL (T)	0	0	1	1	0
NO EMAIL (N)	0	0	0	0	0

Postal mail

	C	S	D	T	N
CARNEGIE MELLON (C)	0	0	0	0	0
STANFORD (S)	0	0	0	0	0
DEVELOPERS (D)	0	0	0	0	0
THEORETICAL (T)	0	0	0	0	0
NO EMAIL (N)	1	0	1	1	1

EMail

	C	S	D	T	N
CARNEGIE MELLON (C)	0	0	1	1	0
STANFORD (S)	0	0	1	1	0
DEVELOPERS (D)	0	1	1	1	0
THEORETICAL (T)	1	1	1	1	0
NO EMAIL (N)	0	0	0	0	0

Group-EMail

	C	S	D	T	N
CARNEGIE MELLON (C)	0	0	1	1	0
STANFORD (S)	0	0	1	0	0
DEVELOPERS (D)	0	0	1	0	0
THEORETICAL (T)	0	0	0	1	0
NO EMAIL (N)	0	0	0	0	0

Figure 3: Cross-Group Communication Mode Usage
NOTE: The internal communication structure of the Soar extended research group was found by running CONCOR on all five communication modes at once. In this figure, each matrix indicates the group-level pattern of communication for a particular communication mode. A 1 indicates that the average group usage of that communication mode to communicate with the other group is greater than the overall average. A 0 indicates that the group average is less than the overall average. In addition, individuals are marked by their current or most recent university affiliation.

TABLE 9
Secondary Information – Invisible College Ties

	Receive No Information by Face-to-Face	*Receive Information by Face-to-Face*	*Total*
Not invisible college	302	166	468
Invisible college	30	152	182
Total	332	318	650
$\chi^2 = 121.01, p < .001$			
	Receive No Information by EMail	*Receive Information by EMail*	*Total*
Not invisible college	366	102	468
Invisible college	60	122	182
Total	426	224	650
$\chi^2 = 118.74, p < .001$			
	Receive No Information by Group-EMail	*Receive Information by Group-EMail*	*Total*
Not invisible college	334	134	468
Invisible college	118	64	182
Total	452	198	650
$\chi^2 = 2.64, p > .1$			

tabulation of actual Group-EMail messages. Most Group-EMail was actually sent by only a few people – developers and the people who maintained the group DLs – and was received by everyone. This mail, largely organizational, was seemingly "forgotten" by the receivers when completing the questionnaire but was more apt to be remembered by the senders. Instead, the members of the group thought more in terms of sending bug reports.[12] In consequence, Figure 3 shows that many people send group mail but few report receiving it.

Ties Used

For the three communication modes (face-to-face, personal EMail, and Group-EMail), we contrasted whether the individual acquired secondary information from another member of this research group through this communication mode with whether the two individuals were related by strong invisible college or proximity ties. There were 650 possible ties (26 × 26, minus ties with self). In determining these ties, the frequency with which an

TABLE 10
Secondary Information—Proximity Ties

	Receive No Information by Face-to-Face	*Receive Information by Face-to-Face*	*Total*
Not proximal	323	207	530
Proximal	9	111	120
Total	332	318	650
$\chi^2 = 111.84, p < .001$			
	Receive No Information by EMail	*Receive Information by EMail*	*Total*
Not proximal	395	135	530
Proximal	31	89	120
Total	426	224	650
$\chi^2 = 102.73, p < .001$			
	Receive No Information by Group-EMail	*Receive Information by Group-EMail*	*Total*
Not proximal	383	147	530
Proximal	69	51	120
Total	452	198	650
$\chi^2 = 10.07, p < .005$			

individual used that mode of communication to interact with the other was considered to be irrelevant (see Tables 9 and 10).

Both face-to-face and personal EMail appeared to support both invisible college and proximity ties (that is, in all cases, we can reject the null hypotheses that the communication mode and the type of tie are independent). In addition, if the individuals received secondary information via either one-on-one, face-to-face interaction or personal EMail, they were almost equally likely to share an invisible college tie with their partner, whereas, they were only one half to two thirds as likely to have a proximity tie. If the individuals did not receive information via these two communication modes, then they were rarely tied by either invisible college or proximity ties. In contrast, Group-EMail and invisible college ties did not appear to be related, and Group-EMail appeared to support proximity ties (although less strongly than personal EMail or face-to-face communication). A somewhat different pattern emerged for Group-EMail. Individuals who received information via Group-EMail were twice as likely not to have invisible college ties and three times as likely not to have proximity ties.

A Note on Collaboration

This study was not about collaboration per se. We did, however, sporadically monitor the collaborative behavior of the members of the Soar extended research group to determine whether new collaborations arose. We found that only those pairs that began collaborating while in the same organization continued this collaboration electronically once they were separated. Individuals who were linked primarily by EMail, even when working on similar projects, did not collaborate. For example, there were potential collaborative pairs between Stanford and Carnegie Mellon graduate students and between Carnegie Mellon and Michigan graduate students, but no collaborations emerged. Instead, each student carefully carved out a separate research area. Clearly, EMail played an important role in facilitating ongoing collaborative efforts, but it did not engender completely new collaborative efforts among strangers.

Our study focused on a single research group that had grown beyond the boundary of a single institution. How EMail will be used when the ties are not as strong as they were within this group is a question for future research. Given the strength of the ties between these researchers, however, the fact that EMail did not promote the development of alternative ties nor collaborations among individuals who were never proximal suggests that EMail will not promote such links between less well-connected individuals.

A Look Toward the Future

Electronic forms of communication can be used to support the basic infrastructure of science, but our study shows that EMail does not stimulate new relationships; instead, it enhances the impact of strong invisible college and proximity ties. EMail has this enhancing effect, in part, because it makes it possible to maintain previous contacts and to extend the research group in terms of number of members. Face-to-face interaction, and a certain level of such interaction, is required initially to establish strong ties. Once strong ties are established, however, EMail can be used to nurture and maintain them.

As electronic mail systems become more prevalent, we might expect to see patterns of communication and diffusion similar to those recorded in this study. For example, as research groups emerge, they may immediately set up a corresponding EMail group, allowing a group to grow arbitrarily large and to be geographically dispersed. Further, as colleagues move among institutions, they will probably use EMail, in addition to other communication modes, to maintain their previous proximal contact and thereby increase the likelihood of extended collaborations that last for more years and involve more people than previously possible with other communication modes. Even new graphic capabilities and expert interfaces that allow presorting of mail, however, are unlikely to result in more diffusion of primary information through EMail unless they are accompanied by technologies for adequate information security and by development of policies for recognizing scientific contributions that are initially communicated electronically.

Telecommunication technologies will continue to be used to communicate secondary information. This pattern was followed in the case of cold fusion. Primary information was announced at a Salt Lake City press conference where the authors staked their knowledge claim. Other scientists gathered the important information they needed to evaluate this claim (secondary information) from *Nature* and the *Journal of Electroanalytical Chemistry*. Telecommunication technologies were used to communicate this secondary information (see Crease & Samios, 1989). As the piece of information grows older and becomes a general part of the literature, then EMail may be used to communicate it to scientists unfamiliar with the idea. And, even with changes in EMail, strong invisible college and proximity ties are unlikely to play a smaller role in the communication of scientific information; instead, changes in EMail will probably serve to enhance the impact of these ties and to maintain their existence and strength. Whether EMail can enhance or maintain invisible college ties that are not based on a mentor-student or co-student relationship is a question for future research.

It is instructive to contrast the findings in this study regarding EMail with those on the telephone. Thorngren (1977), in summarizing the results of multiple studies, noted that telephone contacts "are enriched by elements of face-to-face meetings as part of the contact chain over time. Once one has the background information from earlier meetings, narrower channels (such as the telephone) become more powerful than they would in isolation" (p. 382). Similarly, an implication of our study is that EMail, especially personal EMail, is more powerful if the communicators have met face-to-face. Further, Thorngren argued that "without repeated brief telephone contacts, the personal relationship that is exercised only occasionally might atrophy with time. At the same time, without personal contact as background, telephone networks do not develop" (p. 383). Similarly, an implication of our study is that EMail, like the telephone, helps to maintain relationships even after the communicators have geographically dispersed and that without personal contact, personal EMail networks do not develop and without continued interaction they may falter or fade. This is not to say that people who have not met face-to-face never contact each other initially through EMail,

but such interaction is the exception rather than the rule. Furthermore, such exceptional behavior is not what will characterize telescience.

As Price (1963) pointed out in discussing "big science,"

> the science we have now so vastly exceeds all that has gone before, we have obviously entered a new age that has been swept clear of all but the basic traditions of the old. Not only are the manifestations of modern scientific hardware so monumental that they have been usefully compared with the pyramids of Egypt and the great cathedrals of medieval Europe, but the national expenditure of manpower and money on it have suddenly made science a major segment of our national economy. (p. 2)

Given the cost of big science, among other factors, larger and more interdisciplinary groups are expected to make more progress on a project than each of the individuals working alone. The increased interest of the U.S. National Science Foundation and similar organizations in funding research centers speaks to this belief. If this is the future for science, then even though EMail may not be used for the diffusion of primary information and may not lead to the formation of new collaborative pairs, it will play a vital role in science by making it possible for collaborative pairs to maintain ties and by extending and strengthening invisible colleges.

Notes

1. According to Aborn (1988), the term telescience "is attributed to the National Aeronautics and Space Agency, where it is used to refer to scientific research carried out remotely via computer-based networks, involving collaboration between scientists far removed from each other and often from their physical facilities" (p. 10).

2. For the programs studied, and particularly for the Soar group, co-student ties were stronger than cohort ties. In other graduate programs, cohort ties may be stronger than ties based on having the same mentor.

3. R1 is a backplane configurer for vaxes and PDP computers. It was originally developed for Digital Equipment Corporation by John McDermott (1981) and his colleagues.

4. An extended analysis of all individual messages was not conducted and so information on the range and variation of the data reported in Table 2 is not available. Discussions with Soar group members suggest that the observed pattern is not atypical for normal periods and that during the month preceding a workshop, the flow of Soar-related messages increases.

5. Each message was read for content and then coded by category. For a sample of 30 messages, two coders agreed on how 90% of them should be coded.

6. We first noted this relationship when reading and coding all 682 messages in Table 3. To check our intuition, we took a sample of 40 messages and examined the average text length for messages in each category. Text length was defined in terms of character count.

7. Significance tests are done using a simple difference in means test that compares the use of two particular communication modes, either of which is treated as a binomial variable, and that assumes a 95% level of significance.

8. While we did not study the behavior of these other users, we do know that some of the requests for the Soar code came via EMail.

9. We say that these modes of communication and topic are dominant because discussions with members of the Soar extended research group and our own observation suggested that almost all messages on that topic between those individuals went by that mode of communication. What information people received appeared to be strongly defined by which of these groups they belonged to as members.

10. We feel that the respondents clearly distinguished between these two questions. Not only was different language used by these researchers to describe operationalizations and theory, but the time they reported for first learning the "big ideas" generally preceded the time they reported for first learning of the operationalization.

11. Pairwise difference-in-means tests between chunking and search for each mode of communication showed no difference in the pattern of response. Pairwise difference-in-means tests between chunking and Soar and between search and Soar for each mode of communication showed significant differences for face-to-face and "other oral" modes.

12. This anomaly seems to be the main case in which the respondents interpreted the questionnaire differently than we intended. For other questions, our observations of this group and discussion with them tended to confirm the answers provided to the questionnaire.

References

Aborn, M. (Ed.). (1988). Telescience: Scientific communication in the information age. *Annals of the American Academy of Political and Social Science*, 495.

Aborn, M., & Thaler, A. (1988). On-line research-support systems. *Annals of the American Academy of Political and Social Science, 495*, 127-134.

Allen, T. J. (1977). *Managing the flow of technology: Technology transfer and the dissemination of technological information within Rand organization*. Cambridge: MIT Press.

Anderson, J. G., & Jay, S. J. (1985). The diffusion of medical technology: Social network analysis and policy research. *Social Quarterly, 26*, 49-64.

Barnlund, D. C., & Harland, C. (1963). Propinquity and prestige as determinants of communication networks. *Sociometry, 26*, 466-479.

Bierly, E. (1988). The World Climate Program: Collaboration and communication on a global scale. *Annals of the American Academy of Political and Social Science, 495*, 106-116.

Breiger, R. L., Boorman, S. A., & Arabie, P. (1975). An algorithm for clustering relational data with applications to social network analysis and comparison with multidimensional scaling. *Journal Mathematical Psychology, 12*, 328-383.

Carley, K. M. (1990). Structural constraints on communication: The diffusion of the homomorphic signal analysis technique through scientific fields. *Journal of Mathematical Sociology, 15*, 207-246.

Cole, J. R., & Cole, S. (1973). *Social stratification in science*. Chicago: University of Chicago Press.

Coleman, J. S., Katz, E., & Menzel, H. (1966). *Medical innovation: A diffusion study*. New York: Bobbs-Merrill.

Crane, D. (1970). *Invisible colleges: Diffusion of knowledge in scientific communities*. Chicago: University of Chicago Press.

Crease, R. P., & Samios, N. P. (1989, September 24). Cold fusion confusion. *New York Times Magazine*, pp. 34-38.

Festinger, L. (1950). Informal social communication. *Psychological Review, 57*, 271-282.

Festinger, L., Cartwright, D., Barber, K., Fleichl, J., Gottsdanker, J., Keysen, A., & Leavitt, G. (1948). The study of a rumor: Its origin and spread. *Human Relations, 1*, 464-486.

Freeman, L. C. (1984). Impact of computer-based communication on the social structure of an emerging scientific specialty. *Social Networks, 6*, 201-221.

Frydol, A., & DeGreve, J. P. (1986). Communication activities in scientific disciplines in Belgium. *Scientometrics, 9*(1-2), 37-49.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology, 68*, 1360-1380.

Granovetter, M. S. (1974). *Getting a job: A study of contacts and careers*. Cambridge, MA: Harvard University Press.

Hesse, B. W., Sproull, L., Kiesler, S., & Walsh, J. P. (1990). *Computer network support for science: The case of oceanography*. Working paper, Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA.

Hesse, B. W., Werner, C. M., & Altman, I. (1988). Temporal aspects of computer mediated communication. *Computers in Human Behavior, 4*, 147-165.

Hiltz, S. R. (1978). Impact of a computerized conferencing system upon scientific-research specialties. *Journal of Research Communication Studies, 1*, 111-124.

Hiltz, S. R., & Turoff, M. (1978). *The network nation: Human communication via computer*. Reading, MA: Addison-Wesley.

Huff, C., Sproull, L., & Kiesler, S. (1989). Computer-communication and organizational commitment -- Tracing the relationship in a city government. *Journal of Applied Social Psychology, 19*, 1371-1391.

Johnson, J. C. (1986). Social networks and innovation adoption: A look at Burt's use of structural equivalence. *Social Networks, 8*, 343-364.

Kara-Murza, S. G. (1981). Problems in studying the diffusion of technological innovations in science. *Social Studies of Science, 2*(6), 117-139.

Katz, E., Levin, M. L., & Hamilton, H. (1963). Traditions of research on the diffusion of innovations. *American Sociological Review, 28*, 237-253.

Kerr, E. B., & Hiltz, S. R. (1982). *Computer-mediated communication systems*. New York: Academic Press.

Kiesler, S., Siegel, J., & McGuire, T. (1984). Social psychological aspects of computer-mediated communication. *American Psychologist, 39*, 1123-1134.

Knorr-Cetina, K. (1981). *The manufacture of knowledge*. Oxford, England: Pergamon.

Laird, J. (1986). Soar user's manual: Version 4.0 (Report No. ISL-15). Palo Alto, CA: Palo Alto Research Centers.

Laird, J., Newell, A., & Rosenbloom, P. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence, 33*, 1-64.

Laird, J., Rosenbloom, P., & Newell, A. (1986a). Chunking in Soar: The anatomy of a general learning mechanism. *Machine Learning, 1*, 11-46.

Laird, J., Rosenbloom, P., & Newell, A. (1986b). *Universal subgoaling and chunking*. Boston: Kluwer.

Lievrouw, L. (1986). *The communication network as interpretive environment: "Sense-making" among biomedical research scientists*. Unpublished doctoral dissertation, University of Southern California.

Lievrouw, L. (1988, February). *When a new medium is commonplace: Communication among computer scientists using electronic mail*. Paper presented at the annual meeting of the International Network for Social Network Analysis, San Diego.

McDermott, J. (1981). R1: The formative years. *AI Magazine, 2*, 21-29.

Menzel, H. (1960). *Review of studies in the flow of information among scientists*. New York: Columbia University, Bureau of Applied Social Research.

Nelson, C., & Pollock, D. (Eds). (1970). *Communication among scientists and engineers*. Lexington, MA: D. C. Heath.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Price, D. J. de Solla. (1963). *Little science, big science*. New York: Columbia University Press.

Price, D. J. de Solla. (1965). Networks of scientific papers. *Science, 149*, 510-515.

Price, D. J. de Solla. (1970). Citation measures of hard science, soft science, technology, non-science. In C. Nelson & D. Pollock (Eds.), *Communication among scientists and engineers*. Lexington, MA: D. C. Heath.

Rainwater, L., & Smeeding, T. (1988). The Luxembourg income study: The use of international telecommunications in comparative social research. *Annals of the American Academy of Political and Social Science, 495*, 95-105.

Rice, R. E. (1980). The impacts of computer-mediated organizational and interpersonal communication. In M. Williams (Ed.), *Annual review of information science* (Vol. 15, pp. 221-249). White Plains, NY: Knowledge and Industry Publications.

Rice, R. E. (Ed.). (1984). *The new media: Communication, research, and technology*. Beverly Hill, CA: Sage.

Rice, R. E. (1987). Computer-mediated communication systems and organizational innovation. *Journal of Communication, 37*, 65-94.

Rice, R. E., & Case, D. (1983). Electronic message systems in the university: A description of use and utility. *Journal of Communication, 33*, 131-152.

Rogers, E. M. (1983). *Diffusion of innovations*. New York: Free Press.

Sproull, L., & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science, 32*, 1492-1512.

Steinfeld, C. W. (1986). Computer-mediated communication systems. In M. E. Williams (Ed.), *Annual review of information science and technology* (Vol. 21, pp. 167-203). White Plains, NY: Knowledge Industry Publications.

Thorngen, B. (1977). Silent actors: Communication networks for development. In I. de Sola Pool (Ed.), *The social impact of the telephone* (pp. 374-385). Cambridge: MIT Press.

Waldrop, M. M. (1988a). Toward a unified theory of cognition. *Science, 241*, 27.

Waldrop, M. M. (1988b). Soar: A unified theory of cognition? *Science, 241*, 296.

KATHLEEN CARLEY is Associate Professor of Sociology in the Department of Social and Decision Sciences at Carnegie Mellon University in Pittsburgh, PA.

KIRA WENDT worked on this study while a graduate student at Carnegie Mellon University from 1987 to 1989. She is currently a Systems Analyst at Aion Corporation in Chicago.