HOW DO EXPERIENCED INFORMATION LENS USERS USE RULES?

Wendy E. Mackay
Thomas W. Malone
Kevin Crowston
Ramana Rao
David Rosenblitt
Stuart K. Card

October 1988

CISR WP No. 192
Sloan WP No. 3082-89-MS

Center for Information Systems Research
Sloan School of Management
Massachusetts Institute of Technology

Abstract

The Information Lens system provides electronic mail users with the ability to write rules that automatically sort, select, and filter their messages. This paper describes preliminary results from an eighteen-month investigation of the use of this system at a corporate test site. We report the experiences of 13 voluntary users who have each had at least three months experience with the most recent version of the system. We found that:

1. People without significant computer experience were able to create and use rules effectively.

2. Useful rules can be created based on the fields present in all messages (e.g., searching for distribution lists or one's own name in the address fields or for character strings in the subject field), even without any special templates.

3. Many people used rules to sort messages into folders for storage after reading them, rather than to categorize and prioritize messages before reading them; and

4. People used delete rules primarily to filter out messages from low-priority distribution lists, not to delete personal messages to themselves.

How Do Experienced Information Lens Users Use Rules?

The Information Lens is a prototype intelligent information sharing system designed to help users select, sort and prioritize their electronic mail (Malone et. al., 1987a, Malone et. al., 1987b). While many of the ideas from this system are now being incorporated in other systems, (Borenstein & Thyberg, 1988, Gorry et. al., 1988) there have as yet been no systematic empirical studies of how people actually use such systems.

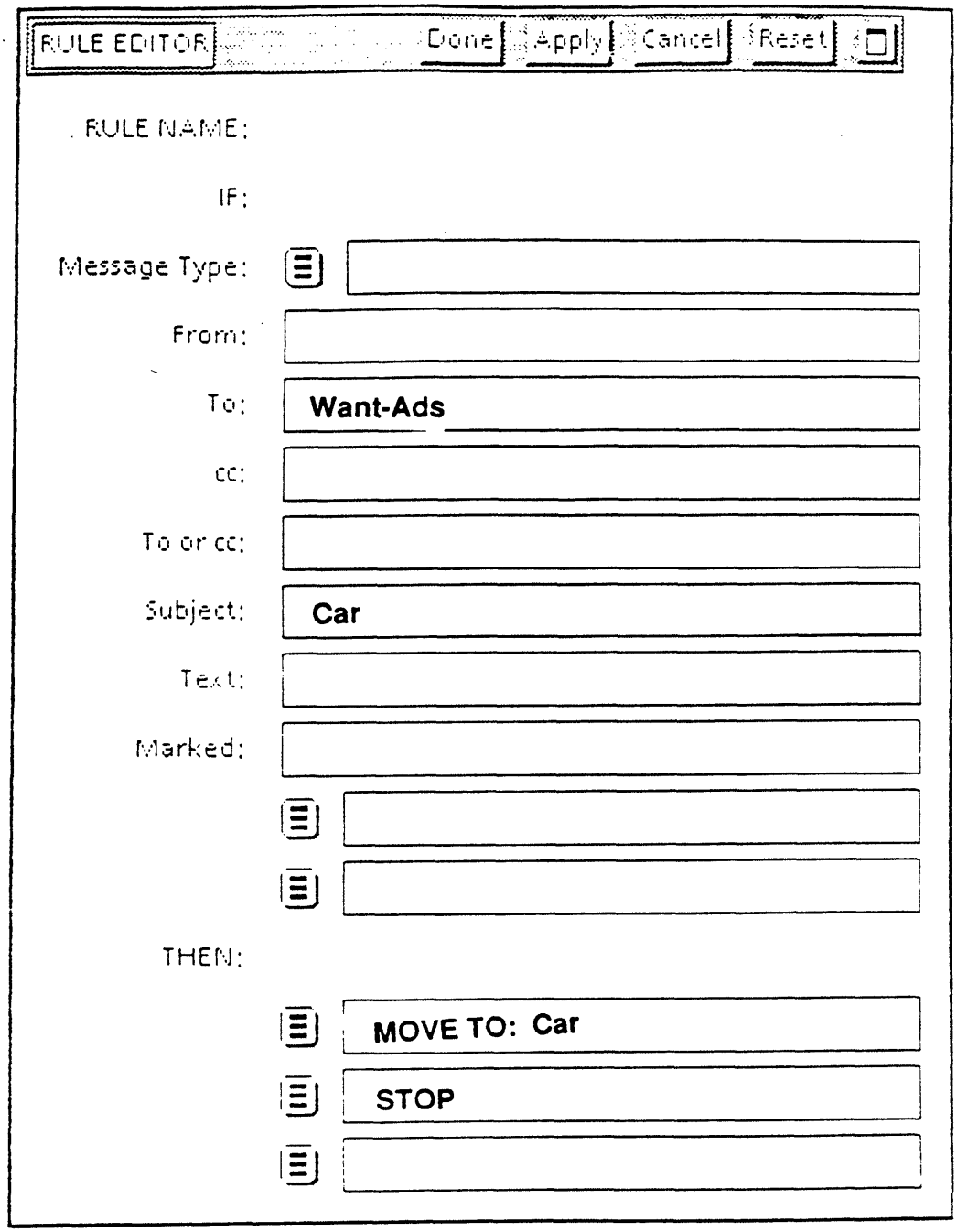
This paper describes the preliminary results of an eighteen-month study of the use of the Information Lens at a corporate test site. The full Information Lens system provides a variety of features, including automatic rules for processing messages, semi-structured message types, and "anyone servers" for routing public messages to interested receivers. This paper concentrates only on the local rules created by experienced users of the system to process messages on their own workstations. Based on repeated interviews with users and on-line records of the rules users created, we have tried to determine how people with significant experience using the system used rules.

We were especially interested in the following questions: Can non-programmers use Lens effectively? What kinds of rules prove useful after several months of experience with the system? When do people choose to run their rules? When do people use rules to delete messages automatically? What kinds of strategies do people use to manage their mail, and how do rules fit these strategies?

Background: Using Rules in the Information Lens System

Users of the Information Lens system can create sets of IF-THEN rules to specify the action they want performed on their messages and the conditions under which these actions should be taken. Common actions include moving the message into a folder, deleting the message, setting a user-defined characteristic (e.g., "Urgent") on the message, and printing the message. Users create rules with a Rule Editor (see figure 1). In this example, the user has specified that messages addressed to the "Want-Ads" distribution list that also contain the subject "Car" are to be moved to the "Car" folder. Whenever a message meets these criteria, Lens will move the message to the "Car" folder. Note that the field labeled "TO or CC" matches text in either field. This feature was added in version 3 in response to many requests; in earlier versions, users had to write two separate rules to achieve this effect.

Figure One


RULE EDITOR
Done
Apply
Cancel
Reset
RULE NAME:
IF:
Message Type:
From:
To:
Want-Ads
cc:
To or cc:
Subject:
Car
Text:
Marked:
THEN:
MOVE TO: Car
STOP

Insert Figure 1 about here

Figure 1. Sample rule that selects messages addressed to the "Want-Ads" distribution list about the subject "Car" and moves them to the "Car" folder.

Users specify when rules should be applied by grouping them into different rulesets. Rules in the "new mail" ruleset are applied automatically when new messages are retrieved to the user's workstation, before the user sees the new messages. Rules in other rulesets ("folder" rulesets) are applied when the user explicitly invokes them on a specified folder. Users primarily invoke folder rulesets after they have at least skimmed the messages in a folder.

Method

Site

The test site for this study was a large laboratory (approximately 60 people) in a research center of a major American corporation. This site was chosen primarily because of (a) its use of the hardware and software environment in which our prototype system was implemented, and (b) its extensive use of electronic mail for a wide range of critical and non-critical organizational communication. The members of this laboratory presumably are more highly educated and more accustomed to advanced computer technology than a representative sample of American office workers. However, as described below, the participants in our study included a spectrum of job categories, educational backgrounds, and amounts of computer experience.

This site has used electronic mail extensively for over a decade and a half. Essentially all employees at this site rely on electronic mail for both formal and informal communication. Users received an average of about 30 to 35 electronic messages per day. Many of these messages come from a variety of distribution lists on both work and non-work topics. The existing mail system in the lab operates in a networked workstation environment that permits the use of separate windows for composing, reading and browsing messages or folders. Messages can be marked with system-defined characteristics, such as moved or

deleted, or with characteristics specified by the user. Users can create and name as many folders as they like for storing messages.

In short, though this site is in some ways unlike many current office environments, it may be fairly typical of sites with lead users of advanced information technology, and we expect that its experiences will resemble those of many other eventual users of advanced electronic communication systems.

Software

The Information Lens prototype has been designed to enhance rather than replace the existing electronic mail system at this site; participants can use as many or as few Lens features as they like. The Lens software is written InterLisp-D and LOOPS and runs on Xerox 1100 series workstations. The first prototype version of the system was introduced to a small group of users in January of 1987. Based on feedback from these users, a second version was created and distributed to a larger group of users in September, 1987. A third version was developed to be used with new versions of the basic mail system and underlying LISP environment. Version 3 incorporated many suggestions from version 2, but omitted other features, including message types and the anyone server (see Malone et. al., 1987). This version became generally available in January 1988, at which time we stopped supporting the earlier versions. Users were only able to convert to version 3 when their workstations were upgraded to the new software. Thus, start dates for the use of version 3 range over a period of several months. Although we have not explicitly encouraged users to use version 3, it continues to be actively used and enhanced at the test site.

Participants

Members of the laboratory were informed about a prototype of the Information Lens and asked if they would be interested in using it. Because we were interested in active users of electronic mail, we solicited participation via electronic mail.

In all three phases of the study combined, over 25 people used some version of the Information Lens software. However, for this paper we are primarily interested in how experienced users create and use rules. Therefore, we will present results from the 13 users who have used version 3 of Lens for at least three months. Some of these users have used this version of the system for as long as seven months

and over half have extensive experience with earlier versions of Lens. The group includes a manager, a manager/secretary team, four computer programmers, one research scientist with formal training in computer science and six other research scientists (from a variety of physical and social sciences) without formal training in computer science. These people used the system in different ways, and we plan to report these differences in a future paper.

Those omitted from this analysis have either stopped using Lens or have not yet accumulated three months of experience with version 3 of Lens. The user who was most heavily involved in implementing the system was also omitted because many of his rules were created to test the system but can not be distinguished from his regular mail processing rules. The primary reason people gave for not continuing use of Lens was conversion to a workstation that does not support the Lens software. Several people were unwilling to try any prototype software. Two people (one programmer and one manager) reported that they did not find Lens sufficiently useful.

In order to help disguise the identities of our subjects in the discussion below, we will refer to all of them using masculine pronouns.

Interviews

Interviews were conducted prior to participation in the study and at two to three month intervals during the course of the study. All interviews were scheduled for one hour in the participant's office. Each person was asked to save the current day's mail and delete confidential messages. Participants were asked to estimate the daily numbers of messages sent and received, the number of mail folders, the size of the in-box and the number of distribution list subscriptions. These answers were checked against the actual numbers for the day and participants were asked if the day was typical. Participants were also asked open-ended questions about major problems and successes with electronic mail. They used this an opportunity to describe their current communication patterns, successful mail management strategies, and problems that needed to be addressed. They also described their use of Lens, including a description of their rules and how they relate to general mail handling strategies. (Refer to Mackay (1988) for a detailed description of the interview format and the actual questions.)

On-line Data Collection

Data about each participant's use of versions 2 and 3 of Lens were collected on-line automatically. Data collected included approximately weekly snapshots of the rulesets and the hierarchy of mail folders and the distribution lists to which each user belonged.

The quantitative data described below are derived from a snapshot of the rules and folders of the 13 most experienced Lens users. For each user, we chose the first day after the user had had three month's experience with version 3 to create a representative, but not exhaustive, sample of rules.

RESULTS

Can non-programming users use rules effectively?

All participants in the study, regardless of computer experience, were able to effectively write their own rules. All users created rules for themselves and none mentioned any difficulty in writing rules. One user, with no computer training, described his first experience with the rule editor as follows: "It's obvious. You just go into the boxes and type whatever you want." However, users do vary in their ability and interest in writing complex rules. Users have the option of writing arbitrarily-complex rules in a LISP-like language, but our interviews indicate most choose to use the graphical rule editor provided with Lens (see Figure 1). This result is consistent with the finding by Jeffries & Rosenberg (1987) that both programmers and non-programmers were able to specify mail filtering rules faster when using a form like those we used than when using a procedural language.

What kinds of rules do people write?

The 13 participants in the study created a total of 180 rules, ranging from 2 to 35 rules per person or an average of approximately 14 each. We classified the rules along two dimensions: the field tested in the IF part of the rule and the action used in the THEN part (see Table 1). Within each field, we have also categorized the most common types of tests in that field (e.g., for a distribution list, for the user him or herself, or for some other person).

Table 1

	ACTION						RULESET				TOTAL	
	Move		Delete		Other		New Mail		Folder			
Recipients	**103**	**57%**	**26**	**14%**	**3**	**2%**	**61**	**34%**	**69**	**38%**	**130**	**72%**
Distribution list	66	37%	20	11%	3	2%	40	22%	47	26%	87	48%
Other string	32	18%	4	2%	0	0%	13	7%	23	13%	36	20%
Self	11	6%	2	1%	0	0%	11	6%	2	1%	13	7%
Other person	4	2%	1	1%	0	0%	4	2%	1	1%	5	3%
From	**22**	**12%**	**10**	**6%**	**0**	**0%**	**22**	**12%**	**10**	**6%**	**32**	**18%**
Other string	9	5%	6	3%	0	0%	12	7%	3	2%	15	8%
Other person	8	4%	2	1%	0	0%	8	4%	2	1%	10	6%
Self	8	4%	2	1%	0	0%	5	3%	5	3%	10	6%
Subject	**46**	**26%**	**14**	**8%**	**0**	**0%**	**30**	**17%**	**30**	**17%**	**60**	**33%**
Project/system	16	9%	4	2%	0	0%	6	3%	14	8%	20	11%
Other string	12	7%	8	4%	0	0%	12	7%	8	4%	20	11%
Distribution list	12	7%	1	1%	0	0%	7	4%	6	3%	13	7%
Meeting/seminar	5	3%	1	1%	0	0%	4	2%	2	1%	6	3%
Phone message	1	1%	0	0%	0	0%	1	1%	0	0%	1	1%
Text or Body	**2**	**1%**	**4**	**2%**	**1**	**1%**	**7**	**4%**	**0**	**0%**	**7**	**4%**
Characteristic	**9**	**5%**	**0**	**0%**	**4**	**2%**	**8**	**4%**	**3**	**2%**	**11**	**6%**
Total	**139**	**77%**	**38**	**21%**	**3**	**2%**	**91**	**51%**	**89**	**49%**	**180**	**100%**

Insert Table 1 about here

Table 1. Bold row headings indicate types of fields. Subheadings indicate types of values within fields. Total rules using a field or field value are shown at the end of each row. Columns 1-6 break down rules by kind of action. Columns 7-10 break down the same rules by kind of ruleset. Total rules using an action or contained in a ruleset are shown at the bottom of each column.

Cell values indicate number and percent of total rules. Columns and rows sum to more than 100% because rules can check more than one field, and (rarely) perform more than one action.

Over half (52%) of the rules involve the processing of distribution lists. Almost all of the participants (85%) use at least one distribution list rule. Several people commented that Lens allows them to stay on distribution lists from which they would otherwise remove themselves. Lens also helps people deal with messages sent to distribution lists to which they belong, not because of their interest in the topic, but because they work in a particular group or want access to a resource controlled by the list.

After distribution lists, the next most common kind of rule identifies messages with some arbitrary string in the RECIPIENTS field (32%). Strings are interesting because they can identify sources of messages in different ways. For example, the string ".BITNET" identifies messages from outside the company, often professional colleagues and friends. Strings can also identify a collection of related distribution lists. For example, the string "AI" selects messages addressed to AI, AI-digest, and AI-interest (as well as, unfortunately, Sailing, etc).

Since the subject is less constrained than other fields, we expected it to be less useful in rules than other fields which usually contain the names of known users or distribution lists. However, a surprisingly high percentage (33%) of the rules select items in the subject field. A breakdown of the kinds of subject fields we found is also given in Table 1.

It is important to note that the total percentage of rules for a particular type is not the only indicator of the importance. For example, each user can have an arbitrarily large number of distribution lists rules, but at most three rules that refer to him or herself, i.e. "TO: user", "CC: user", and "FROM: user." The latter make up only 13% of the total number of rules, yet 85% of the users had a least one and most people described "self" rules as being extremely useful.

Similarly, it is tempting but not accurate to believe that users with more rules find Lens more useful. In fact, some of the users who reported the most satisfaction with Lens had only two to four rules.

When do people run rules?

As mentioned above, rules can be run either when new mail is received (newmail rules) or on demand (folder rules). In our original design of the Information Lens system, we anticipated that people would run rules to sort and prioritize messages before reading them (i.e., when new mail is received). In our initial interviews, however, Mackay (1988) found that a number of users said they would want to use rules to sort messages after reading them. She called the first kind of users "prioritizers" and the second kind "archivers." In this study we confirmed that experienced users did, in fact, choose to use rules at both times. However, individual users often have strong preferences for one or the other. In our sample, five people use newmail rules exclusively ("prioritizers"), three people use folder rules exclusively ("archivers"), and five people use both kinds of rules. An equal number of the total rules were in newmail (51%) and folder rulesets (49%).

What kinds of delete rules do people use?

One of the most commonly asked questions about Lens is what rules people use to delete messages. Approximately one fifth (21%) of the total number of rules delete messages (the breakdown of rules by action is given in Table 1). Of these, the most common use is to delete messages from low-priority distribution lists (74%).

Only five delete rules (13%) refer to a person's name. In each case, the rules use multiple fields to further qualify the type of message. In two cases, the messages deleted are announcements of events sent by an outsider or part-time employee to the entire lab. The senders of these messages are misusing the system, but appear to be outside the usual social pressures that would normally prevent this behavior. Even though it is unlikely that the Lens user will want any particular messages from either sender, the rules are still qualified to filter out only messages about the unwanted subject. Personal messages from either person would not be deleted.

In general, deletion rules appear to be more complex than rules that move messages to folders. If we define complex rules as those with more than one field (counting "TO: or CC:" as a single field), then 60%

of delete rules are complex, whereas only 32% of move rules are complex. One explanation for this is that deleting is more "dangerous" than moving because an error has greater consequences, i.e., the accidental deletion of a message.

What strategies do people use to manage their mail?

[Note to reviewers: If length constraints are binding, this section can be omitted.]

Rules reflect the strategies users have for managing their inboxes. We found people using three primary strategies for handling the new mail in their inboxes:

1. ***Keep it all.*** Process mail as it arrives and leave it in the inbox until it has been successfully handled, regardless of priority (use folder rules or no rules).
2. ***Move the unimportant messages.*** Move inessential messages out of the inbox and use the inbox as a repository for things to do and unprocessed mail (use newmail rules and possibly folder rules).
3. ***Move the important messages.*** Move all important messages into separate high-priority folders and use the inbox as a temporary storage place for lower-priority items (use newmail rules and possibly folder rules).

The most interesting uses of Lens rules occurred in the second and third strategies.

Move the unimportant messages. In this strategy, rules can be used to identify low priority messages. However, several individuals reported that they moved messages from low priority distribution lists into folders only to discover that they never looked at them. People have different reactions to this. Some decide that this provides useful information about their priorities; it's a good technique for rating the usefulness of different distribution lists. Other users do not mind letting the messages pile up and are happy to avoid the clutter in their inboxes.

Another strategy for handling low priority mail is to use newmail rules to automatically delete them. Since the deleted messages appear on the screen with a line drawn through them, the user has a chance to see them before they are expunged from the system. ("I don't remove myself from mailing lists, I just automatically delete the message. I occasionally spot something in the deleted messages and take a look. I'm almost always disappointed and think that, yes, I should have deleted it.") This person made the point that it is easier to write a rule than get on or off a distribution list. He changes his rules from delete to move, depending on his current work load and whether the topic is currently relevant. "It makes the cost of staying on a distribution list very low." One person uses a similar strategy for identifying

potential rules for handling messages in his inbox. "If a rule doesn't fire on a set of messages in the inbox, I let them accumulate and use [those messages] to determine whether a special rule is needed."

Another rule considered useful by a subset of users is the "boring" rule. An early Lens user observed that some distribution lists often carry boring discussions about a particular subject. Messages in a particular discussion can be easily identified, because the subject of the following message usually contains the subject of the original message. Lens version 3 allows a user to point to such a message and automatically create a rule that deletes all subsequent related messages. (This is an example of "rule-creation-by-example.") Those who use this rule extensively subscribe to an above-average number of distribution lists. The four people who use this rule are very happy with it. ("This boring business is a real win!" "This kind of thing is great!") Another user tried it for a while and did not find it particularly useful. This person, however, rarely reads distribution lists.

Move the important messages. In this strategy, rules can be used to identify and move important messages to separate folders. A special case of this is the use of "self" rules to move personally-addressed messages to a priority folder. Five people use newmail rules for this. Each is willing to make the inbox lower priority and discipline themselves to read priority folders first.

A typical example is a senior researcher who stays off all voluntary distribution lists and relies on colleagues to provide him with technical information. He uses mail to maintain an active correspondence with colleagues around the world; these make up two-thirds of his messages. His rules correspond closely to his folders, identifying messages addressed to him personally and messages sent via "BITNET". He has very few rules, but they make a major difference in his use of electronic mail. "When I've been gone, Lens is fantastic. I was away for a week and it was a pleasure to have my mail sorted correctly."

Another user converted to this strategy after using Lens for many months. He was initially very distrustful of Lens, and first decided to use it to help sort previously-skimmed messages in his inbox. He refused to use new mail rules, explaining the "Nothing happens automatically - I look, then sort." This user allowed his inbox to reach an unmanageable state with over 1100 messages. The best he could do was collect the "old" messages from the beginning of the inbox, move them into a dated folder, and hope that there were no important messages within.

This person tried a new strategy, almost by accident, just before converting to version 3. He was preparing to go away on a trip, and rather than converting all of his old rules, he created two new ones. The first identifies all personally-addressed messages and places them into a "priority" folder. The second identifies all messages related to a conference he was running and sends them to the conference administrator. When he returned, he found that this strategy was "very, very useful" and continued using these two rules. He has stopped trying to keep on top of his inbox and instead, reads only the "priority" folder on a regular basis. He reads whatever is left in the inbox whenever he gets a chance and either files or deletes the messages manually. He intends to use Lens for archiving rules: "Once I get organized, it will be extremely useful. I hate to stop long enough to get it set up." After seven months, he still hasn't created folder rules. He said these two rules "changed my life" and no longer describes his mail as being out of control.

Conclusions

The designers of the original Information Lens system anticipated that systems like this would help users to automatically sort and filter their messages before reading them, based on the contents of fields in semi-structured messages templates. People who hear the system described often wonder (a) whether non-programmers will be able to create their own rules, (b) whether useful rules can be created without everyone in a group using the same message templates, and (c) how often people would use the system to automatically delete personal messages from other individuals. In this preliminary study of the system, some of our expectations were confirmed and others were not. We found that:

1. People without significant computer experience were able to create and use rules effectively.

2. Useful rules can be created using only the fields present in all messages (e.g. TO, FROM, CC, and SUBJECT), even without any special templates.

3. Many people used rules to sort messages into folders for storage *after* reading them, rather than to categorize and prioritize messages before reading them; and

4. People used delete rules primarily to filter messages from low-priority distribution lists.

We expect to report further results from this study in future papers, and we hope other studies will eventually examine these same questions in larger scale implementations of similar systems. In the meantime, however, we believe these results provide the first systematic empirical evaluation of people using rule based tools to help manage their personal electronic communication.

Acknowledgements

We would like to thank Maya Bernstein for her cheerful and extensive tabulation of the data. We would also like to thank the participants in the study for their thoughtful comments and willingness to discuss their different strategies for managing electronic mail.

References

Borenstein, N.S. & Thyberg, C.A. (September 1988). Cooperative Work in the Andrew Message System. *Proceedings on the Conference for Computer-Supported Cooperative Work.* Portland, Oregon.

Gorry, G.A., Burger, A.M., Chaney, R.J., Long, K.B., Tausk, C.M. . (September 1988). Computer Support for Biomedical Work Groups. *Proceedings on the Conference for Computer-Supported Cooperative Work.* Portland, Oregon.

Jeffries,R and Rosenberg, J. . (April 1987). Form-based and a Language-Based User Interface for Instructing a Mail Program. *CHI + GI Conference Proceedings.* Toronto, Canada.

Mackay, W.E. (September 1988). More Than Just A Communication System: Diversity in The Use Of Electronic Mail. *Proceedings on the Conference for Computer-Supported Cooperative Work.* Portland, Oregon.

Malone, T.W., Grant, K.R., Lai, K.Y., Rao, R. & Rosenblitt, D.R. (1987). Semistructured Messages Are Surprisingly Useful for Computer-Supported Coordination. *ACM Transactions on Office Information Systems, 5*(2), 115-131.

Malone, T.W., Grant, K.R., Turbak, R.A., Brobst, S.A., & Cohen, M.D. (1987). Intelligent Information-Sharing Systems. *Communications of the ACM, 30*, 484-497.