

A Study of the Use of Electronic Mail in Academic Communications

by
Jonna J. K. Peat

Submitted to the Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degrees of
Bachelor of Science in Computer Science and Engineering
and Master of Engineering in Electrical Engineering and Computer Science
at the Massachusetts Institute of Technology

May 22, 1998

[June 1998]

© Copyright 1998 Jonna J. K. Peat

The author hereby grants to M.I.T. permission to reproduce and
distribute publicly paper and electronic copies of this thesis
and to grant others the right to do so.

Author _____
Department of Electrical Engineering and Computer Science
May 22, 1998

Certified by _____
Thomas J. Allen
Thesis Supervisor

Accepted by _____
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

JUL 14 1998

LIBRARIES

ENE

A Study of the Use of Electronic Mail in Academic Communications

by
Jonna J. K. Peat

Submitted to the
Department of Electrical Engineering and Computer Science

May 22, 1998

in Partial Fulfillment of the Requirements for the Degrees of
Bachelor of Science in Computer Science and Engineering
and Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

The purpose of this research project was to study the use of electronic mail in academic communication. Specifically of interest were how people communicate with e-mail and how they manage their e-mail. The research was carried out as an electronic survey on the world-wide web for which the sample group consisted of members of the Sloan Management School at the Massachusetts Institute of Technology, and the results were analyzed with statistical tests. The data from the survey responses supported seven of the eight original hypotheses and demonstrated that the volume of e-mail one receives has an effect on how carefully one reads and manages their e-mail as well as one's perception of the efficiency of the medium as a communication tool.

Thesis Supervisor: Thomas J. Allen
Title: Deputy Dean, Sloan School of Management

Table of Contents

List of Tables	5
1.0 Introduction	6
2.0 Background	7
2.1 The Definition of Electronic Mail	7
2.2 Relevant Issues	7
2.2.1 Ambiguity	8
2.2.2 A Different Message	9
2.2.3 Efficiency	9
2.2.4 A Larger Audience	10
3.0 Thesis Statement	11
4.0 Hypotheses	11
5.0 Research Design	14
5.1 Preliminary investigation	14
5.2 Sampling	14
5.3 The Survey	15
6.0 Observations and Analysis	17
6.1 Response rate	17
6.2 Statistical tests	17
6.2.1 Bivariate correlations	18
6.2.2 Difference in proportions	19
6.3 Corresponding with e-mail	19
6.3.1 Hypothesis 2: Location of correspondent	20
6.3.2 Hypothesis 3: Affiliation of correspondent	20
6.3.3 Intent of last correspondence	22
6.3.4 Hypothesis 1: E-mail v. face-to-face meetings	24
6.4 Messages received within a day	25
6.4.1 Analysis of hypotheses 4 through 7	26
6.4.2 Hypothesis 4: Messages read thoroughly	26
6.4.3 Hypothesis 5: Messages replied to immediately	27
6.4.4 Hypothesis 6: Deleted messages	27

6.4.5 Hypothesis 7: Useful messages	28
6.4.6 Junk mail	28
6.5 Inbox composition	29
7.0 Results	31
8.0 Commentary	32
8.1 Problems encountered	32
8.1.1 Ambiguity	32
8.1.2 The limitations of multiple-choice	32
8.1.3 Interval questions	33
8.1.4 Differences in e-mail software	34
8.2 Possible improvements	34
8.3 Recommendations	35
9.0 Conclusion	36
10.0 Bibliography	37
11.0 Appendices	39
11.1 The survey	39
11.2 Data tables	42
11.2.1 Hypothesis 1	42
11.2.2 Hypothesis 2	42
11.2.3 Hypothesis 3	43
11.2.4 Hypotheses 4-7	44
11.2.5 Hypothesis 8	44
11.3 Comments from the respondents	45

List of Tables

TABLE 1. Hypotheses	12
TABLE 2. Summary of survey response rates	17
TABLE 3. Affiliation: Significant differences	21
TABLE 4. Message intent: Significant differences	23
TABLE 5. Data conversion examples	26
TABLE 6. Comparison of analyses with and without “20+” data points	28
TABLE 7. Last e-mail v. face-to-face meeting	42
TABLE 8. Frequencies: Location of correspondent.	42
TABLE 9. Proportions: Location of correspondent	43
TABLE 10. Frequencies: Affiliation of correspondent.	43
TABLE 11. Proportions: Affiliation of correspondent	43
TABLE 12. Management of messages received yesterday (Spearman’s <i>r</i>)	44
TABLE 13. Types of messages in inbox (Spearman’s <i>r</i>)	44

1.0 Introduction

In many academic, professional, and personal settings around the world, the popularity of electronic mail has increased dramatically as people try to capitalize on its speed and cost-effectiveness. As its use has increased, however, problems have arisen related to the management of electronic mail. Because it is relatively new, e-mail as a medium suffers from a lack of structure that can result in information overload or general miscommunication. Unlike most other means of communication, the recipient has very little if any control on the volume of information he receives. Given that some subset of this information is important to one's work, it is important that one is able to correctly discern what is relevant and what is not.

If e-mail is to continue to be an effective communications medium, it is important to consider what factors affect whether electronic messages are useful information or information overload. This thesis will explore various issues related to this topic, among them:

1. how e-mail supplements or replaces oral communication (such as face-to-face meetings or telephone conversations),
2. the effect of the volume of e-mail one receives on how carefully one reads and responds to messages,
3. how people manage their e-mail they so as to minimize the time they spend reading irrelevant information,
4. the effectiveness of electronic mailing lists as a communication tool and its relation to the perception of e-mail as useful information, and
5. the types of information people choose to communicate via e-mail.

2.0 Background

2.1 The Definition of Electronic Mail

For the purposes of this report, an electronic mail system is any system that:

1. permits the asynchronous electronic interchange of readable information between persons, groups of persons, and functional units of an organization; and
2. provides mechanisms supporting the creation, distribution, consumption, processing, and storage of this information (Anderson, 1989).

It should be noted that by this definition, any messages or files exchanged over a computer network qualify as electronic mail whereas other forms of electronic communication such as faxes or voice mail do not.

The specification that e-mail is asynchronous implies that electronic messages can be delivered to someone's mailbox without the recipient's being physically present or even logged on to the network. This asynchrony is an important factor in the power of e-mail as a communications tool. To illustrate, up to 70% of telephone calls are never completed because the receiving party is not available (Philip and Young 1987 quoted in Sproull, 1991).

2.2 Relevant Issues

There are many advantages to the use of electronic mail over other standard means of communication. Face-to-face meetings, telephone conversations, written letters, etc. all differ on the speed of initiating contact and transmitting information, the cost of distri-

bution, the permanence of the message, accessibility, privacy, and their accountability and attributability (Shapiro/Anderson, 1985). Electronic mail is faster, cheaper and often more reliable as a means of communication; however, along with these benefits come several detrimental side effects. The sections below discuss some of the factors that undermine the effectiveness of e-mail as a means of communication.

2.2.1 Ambiguity

One important disadvantage of this medium is that it is impoverished in the social cues to which we are so accustomed. In other established media, social cues are an important part of what is being communicated. A study done in the 1970s concluded that 93% of peoples' intent is conveyed by their tone of voice and facial expression while speaking (Meherabian 1971 quoted in Anderson, 1989). With telephone conversations or face-to-face meetings, for example, the participants are able to utilize body language to formulate their thoughts about the content of the verbal message. E-mail does very little to convey humor, sarcasm, nuance, or other deeper meaning.

Body language influences the speaker as well. People choose the words they use based on the context of the conversation. If the listener appears confused, for instance, the speaker can immediately resolve any ambiguities. If the listener seems to be getting angry, the speaker can curb his/her expressions to make them less blunt or offensive. E-mail does not offer this immediate feedback which can sometimes result in miscommunication.

Because e-mail is relatively new, there is no instruction or experience to guide

what effective communication is as there is with other media. At least one attempt has been made to remedy this by the publication of the handbook *Toward an Ethics and Etiquette for Electronic Mail* in the mid-1980s. This book suggested practical guidelines for sending and receiving electronic mail. The goal was to counteract some of the negative aspects of this quick and easy medium where reading between the lines can be the source of serious problems. Despite these efforts, these issues continue to plague the effectiveness of e-mail as a communication tool.

2.2.2 A Different Message

Proponents of computer-based communications say it delivers the same message as oral communication only more concisely and rapidly; however, research in the field has provided a lot of evidence to the contrary. In a survey of alcohol consumption, for example, the data collected in an electronic survey was much closer to reported alcohol sales than data from the paper-based survey. Other experiments show that people are more likely to exaggerate their good points and downplay their bad points in physically-based communications than electronically-based communications (Sproull, 1991). Though this seems to imply that electronic mail promotes honesty, it is important to consider how and why people change their messages-- and whether e-mail could actually have a totally opposite effect under certain circumstances.

2.2.3 Efficiency

Studies show that the speed and effectiveness of communications in a group determine the efficiency with which they work together (Aldredge, 1980). In this regard, e-

mail is a boon to productivity; however, more efficiency in one place often means less in another. A company might give its employees e-mail access which increases the speed of their communications and the number of people they can communicate with, but it also creates increased overhead in the time spent training employees to utilize the electronic mail system as well as the time spent checking and responding to e-mail.

2.2.4 A Larger Audience

One of the societal implications of e-mail is that it changes how people spend their time and who they know and care about (Sproull, 1991). For example, an employee searching for the answer to an important question may broadcast the message to everyone in the company. This is good in that the employee does not have to know exactly what resources there are for the information (as one would have to in order to initiate person-to-person communication such as a telephone call) to resolve the problem. Also, he/she does not waste time contacting people who do not know the information in order to be forwarded to the correct sources. Those who can answer the question can respond directly to the employee's message without having to know who he/she is.

On the other hand, such messages incur a lot of wasted time. Though it may be true that 50 of the company's 5,000 employees are able to answer the question, this message had to be read and/or deleted by 4,950 other employees. In any business, time is money, so the efficiency of broadcasting such a message may not be worth the time (and possible annoyance) of all the people forced to read it. As this example illustrates, more information is not always better than less information.

3.0 Thesis Statement

The purpose of this thesis is to study the use of e-mail as a communication tool in an academic setting. By determining how people use e-mail to communicate with others and how they manage the e-mail they receive, I hope to shed light on a medium that as yet still has relatively few protocols in place. This information may produce beneficial results for those implementing policies on the use of electronic mail as a medium for communication as well as those seeking to implement communications software that is both useful and effective.

4.0 Hypotheses

The survey will be designed to test hypotheses related to the effectiveness of e-mail as a communications tool within an academic setting. Based on my research, I make the following hypotheses:

1. The degree to which people choose to communicate by e-mail is dependent on how often they see each other face-to-face.
2. People are just as likely to use e-mail to communicate with those people located far away as they are to communicate with those people close by.
3. People are more likely to e-mail those within their work group than people outside of their work group.
4. The larger the volume of e-mail one receives, the less likely one is to read those messages thoroughly.

5. The larger the volume of e-mail one receives, the less likely one is to reply immediately to those messages that require it.
6. The larger the volume of e-mail one receives, the more likely one is to delete messages without reading them.
7. The larger the volume of e-mail one receives, the less likely one is to find those messages useful.
8. A person's position/role within an organization affects how they use e-mail as a communications tool.

TABLE 1 presents a nomological net for each of the hypotheses. It summarizes the independent and dependent variables as well as the related data that would test the validity of the hypotheses.

TABLE 1. Hypotheses

	Independent Variables	Dependent Variables
Hypothesis 1: The degree to which people choose to communicate by e-mail depends on how often they see each other face-to-face.	The choice of e-mail as a communication tool.	The frequency of face-to-face meetings between the correspondents.
RELATED DATA	The last person the participant sent e-mail to.	The last time the correspondents met face-to-face.
Hypothesis 2: People choose to use e-mail to communicate regardless of the location of the correspondent.	The choice of e-mail as a communication tool.	The location of the correspondents.
RELATED DATA	The last person the participant sent e-mail to.	Their relative location to the correspondent.
Hypothesis 3: People are more likely to e-mail people closely affiliated to them.	The choice of e-mail as a communication tool.	Belonging to the same work group.
RELATED DATA	The last person the participant sent e-mail to.	The smallest work group (if any) both correspondents belong to.

TABLE 1. Hypotheses

	Independent Variables	Dependent Variables
Hypothesis 4: People are less likely to read their messages thoroughly if they receive a large volume of e-mail.	The volume of e-mail received	What the participant reads
RELATED DATA	The number of messages the participant received the day before.	The number of messages the participant read all the way through.
Hypothesis 5: The larger the volume of e-mail one receives, the fewer messages one replies to immediately.	The volume of e-mail received	What the participant replies to
RELATED DATA	The number of messages the participant received the day before.	The number of messages the participant replied to immediately.
Hypothesis 6: The larger the volume of e-mail one receives, the more messages one deletes without reading them.	The volume of e-mail received	What the participant deletes
RELATED DATA	The number of messages the participant received the day before.	The number of messages the participant deleted the day before.
Hypothesis 7: The larger the volume of e-mail one receives, the less likely one is to find those messages useful.	The volume of e-mail received	How many messages the participant finds useful.
RELATED DATA	The number of messages the participant received the day before.	The number of messages the participant found useful.
Hypothesis 8: A person's position/role within an organization affects how they use e-mail as a communications tool.	The participant's role within the organization.	How the participant uses e-mail to communicate with others: The frequency of face-to-face meetings with e-mail correspondents; the location of the correspondents; belonging to the same work group. How the participant manages their e-mail: what mail the participant reads and replies to.
RELATED DATA	The participant's position within the Sloan School	<i>How the participant uses e-mail to communicate with others:</i> the last time the correspondents met face-to-face; the relative location of the correspondents; the smallest work group both correspondents belong to. <i>How the participant manages their e-mail:</i> the number of messages he/she read all the way through, deleted, responded to immediately, found useful to his/her work.

5.0 Research Design

5.1 Preliminary investigation

The research for this project was to be conducted as a survey on the world-wide web. After determining the hypotheses, questions were formulated to test them appropriately. Once the first draft of the survey was completed, a group of eight students were asked to complete it. The purpose of this pre-survey was to:

- determine that all the electronic features of the survey (multiple-choice lists, submission to the web server, etc.) worked on different operating systems,
- correct any ambiguities,
- test which format (multiple-choice, open-ended) worked best for each question,
- discover any additional options to be included on multiple-choice questions, and
- improve the general readability of the survey.

The comments of all the pre-survey participants were taken into account in the final revision of the survey.

5.2 Sampling

The intended audience for the survey was all members of the Sloan School at MIT. An e-mail was sent out to the Sloan School mailing list directing interested volunteers to the URL¹ where the survey was located. Respondents were able to point and click on

1. <http://allen-wrench.mit.edu/survey/fac-email.html>

responses to the survey and submit them electronically to a central database. All participants were advised that the survey was completely anonymous and voluntary.

The Sloan School mailing list is comprised of about 1,375 people each of which has one of the following titles:

- **faculty** (126): professors
- **research staff** (21): employees who do research but do not teach
- **support staff** (61): secretaries
- **administrative staff** (67): responsible for department programs, centers, and other administrative tasks (e.g. accounting)
- **students** (1,100): all students registered in the Sloan School.

5.3 The Survey

The survey¹ was divided into several sections intended to solicit responses that would appropriately test the hypotheses. The first question determined the respondent's title in the Sloan School. The next section asked questions about the e-mail the participant received during the course of a day. Specific issues addressed were the volume of mail, how often they check for new mail and how much of that mail was "junk mail."

The following section asked questions about the first batch of e-mail the participant read that day. These questions were designed to determine the relative composition of those messages-- how many they read all the way through, how many they deleted

1. A copy of the survey can be found in Section 11.1.

before reading, how many they responded to immediately, and how many provided useful information.

The next section asked questions about the messages currently in the participant's inbox. I made the assumption that the messages people keep in their inbox are the ones they want to pay attention to. With this in mind, the questions in this section determined the types of messages the respondent chose to keep.

The last section was designed to test how the participants use e-mail to communicate with others. It asked questions about the last person (not group) to whom they sent a message-- where the correspondent was located, their affiliation, the last time they spoke face-to-face, etc. Finally the participants were asked to supply any comments they had about the survey and its content.

6.0 Observations and Analysis

6.1 Response rate

This survey method turned out to be largely successful. Out of the 1,375 people on the Sloan School mailing list, 233 people responded within a three-day period. Of the students 150 out of 1,100 filled out the survey-- a 13.6% response rate. Of the Sloan School employees (faculty and staff), 83 out of 375 replied-- a 22.1% response rate. The survey was short; it had a total of 19 questions, 10 of which were multiple choice. The other nine were quantitative questions for which the participant could easily determine the answers (e.g. "Of the messages currently in your inbox, how many are meeting reminders/ requests?"). Another factor that may have helped is that there was very little effort associated with turning the survey in since all it required was the click of a button.

TABLE 2. Summary of survey response rates.

Sloan School Title	# in Sloan	# of Survey Respondents	Response Rate
faculty	126	31	24.6%
research staff	21	6	28.6%
support staff	61	19	31.1%
administrative staff	67	27	40.3%
students	1,100	150	13.6%
TOTAL	1,375	233	16.9%

6.2 Statistical tests

Three different statistical tests were used to analyze the data. They are summarized below. All data was analyzed for significance at the 0.01 level.¹

1. A significance level of .01 means that there is only a 1% chance that our conclusion about whether or not to reject the hypothesis is incorrect (Dixon and Massey, 1957).

6.2.1 Bivariate correlations

- *Pearson's correlation coefficient*

Pearson's correlation coefficient (r) is used to measure the strength of the linear relationship between two normally distributed quantitative variables (SPSS Inc., 1996). The value of r falls between -1.0 (which indicates an inverse relationship) and +1.0 (which indicates a direct relationship). If the correlation coefficient is close to zero, it indicates that the two variables are independent of each other. The formula is:

$$r = \sum_{i=1}^N \frac{(x_i - \bar{x})(y_i - \bar{y})}{(N-1)s_x s_y}$$

(EQ 1)

In this formula, N is the size of the sample, \bar{x} and \bar{y} are the means of the two variables, and x_i and y_i are the actual values recorded for case i .

- *Spearman's rank correlation coefficient*

The Spearman rank correlation coefficient is also used to measure the strength of the linear relationship between two variables, but there is no assumption that the variables are normally distributed and/or quantitative (SPSS Inc., 1996). Spearman's test uses the same formula as Pearson's (EQ 1), but it uses the rank order of each data value for the variable. In the case of ties, adjustments are made. This statistic also varies between -1.0 and +1.0.

6.2.2 Difference in proportions

The test for difference in proportions determines whether or not there are significant differences in the proportions of two populations having a certain characteristic (Dixon and Massey, 1957). Its formula is given below.

(EQ 2)

$$\bar{X}_1 - \bar{X}_2 + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\bar{X}_1(1-\bar{X}_1)}{N_1} + \frac{\bar{X}_2(1-\bar{X}_2)}{N_2}} < p_1 - p_2 < \bar{X}_1 - \bar{X}_2 + z_{\frac{\alpha}{2}} \sqrt{\frac{\bar{X}_1(1-\bar{X}_1)}{N_1} + \frac{\bar{X}_2(1-\bar{X}_2)}{N_2}}$$

In EQ 2, N is the size of the sample, \bar{X} is the number in the sample having the characteristic / N , α is the level of significance, and p is the proportion of the population having the characteristic.

If the left side of the equation is negative and the right side is positive (e.g. $-0.234 < p_1 - p_2 < -0.416$), then there is no reason to doubt the null hypothesis that $p_1 = p_2$, that is, there is no significant difference between the proportions of the sample populations exhibiting the characteristic. However, if both sides have the same sign, then the null hypothesis should be rejected.

6.3 Corresponding with e-mail

The survey asked several questions about the last person to whom the respondent composed a message. There are two subtle distinctions here which should be emphasized.

1. Any messages that were sent out to a group mailing list were not to be included.
2. Replies to messages were not to be included.

Below are the analyses for the hypotheses related to messages sent by the participants.

6.3.1 Hypothesis 2: Location of correspondent

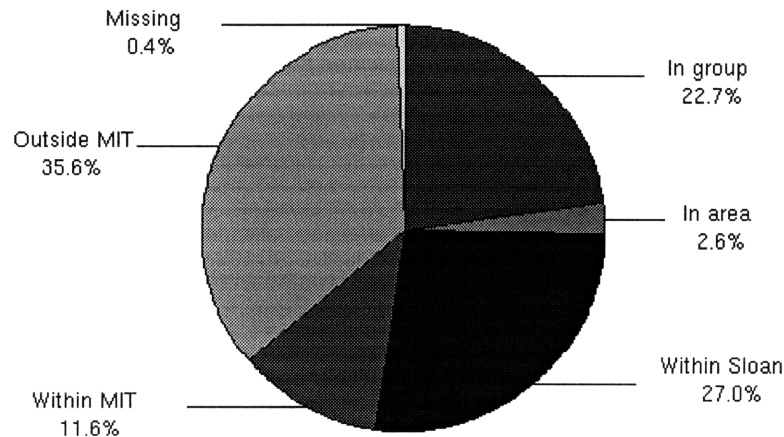
The first question in this section asked for the location of the last person to whom the respondent sent e-mail. The choices were: same floor, same building, different building, and outside MIT.

The data was analyzed using Spearman's bivariate correlation to see if there was a relationship between location and the number of messages composed. This test revealed a perfect correlation (1.0) in the aggregate-- that is, more e-mails are composed to people farther away. This seems to refute the original hypothesis that people are equally likely to communicate with those near by as with those far away; however, the farther away the category, the greater the number of people in it (i.e. there are more people "outside MIT" than on the "same floor"). In the end, the data does not support the rejection of the null hypothesis, namely, who people choose to communicate with via e-mail is not related to their relative location.

6.3.2 Hypothesis 3: Affiliation of correspondent

The next question asked for the closest affiliation of the correspondent to whom the participant last sent e-mail. The choices were: in group, in area, within Sloan, within MIT, and outside MIT. The data (summarized below in FIGURE 1) was analyzed using the test for differences in proportions (at the .01 level of significance) to detect whether or not one's position within the Sloan School was correlated with this variable.

FIGURE 1. Affiliation of correspondent



The results show a few significant differences in the proportions of people in each position who communicate within Sloan and outside MIT (supporting Hypothesis 3). These are illustrated in **TABLE 3**. Other than these, there are no significant differences that would illustrate that people in different positions in the Sloan School communicate with people with any particular affiliation.

TABLE 3. Affiliation: Significant differences

Affiliation	Position 1	Proportion of messages (1)	Position 2	Proportion of messages (2)	Difference in proportion
Within Sloan	Student	32%	Research staff	0%	$-0.417 < p_1 - p_2 < -0.222$
Outside MIT	Faculty	53%	Administrative staff	15%	$-0.092 < p_1 - p_2 < -0.667$
Outside MIT	Students	37%	Administrative staff	15%	$-0.420 < p_1 - p_2 < -0.010$

Though the statistical test shows that there are relatively few significant differences, one interesting observation about the affiliation of the last person to whom the respondent sent e-mail was that for each group (except for the support staff) there was one particular type of affiliation with which they corresponded more than any other group. For administrative information, it was people within their group (33%); for the faculty, people

outside MIT (54%); for the research staff, people within MIT (33%); for the students, people within Sloan (32%).

6.3.3 Intent of last composed message

The third question asked the respondents to pick from the following list the response that most closely fit the intent of their last composed message: asking a question, sending files, setting up a meeting, sharing an idea/opinion, other. The data was analyzed using the test for differences in proportions (at the .01 level of significance) to see whether or not different groups of people within the Sloan School used e-mail for different purposes.

FIGURE 2. Intent of last composed message, non-students.

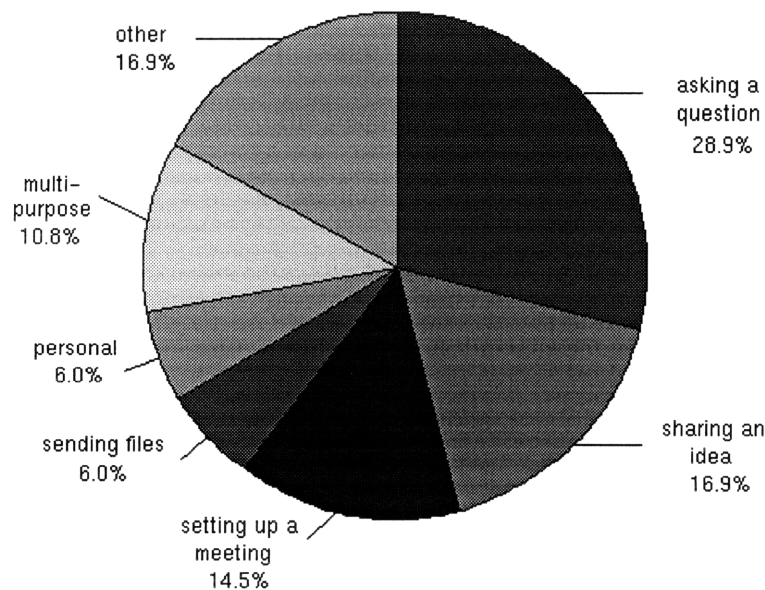
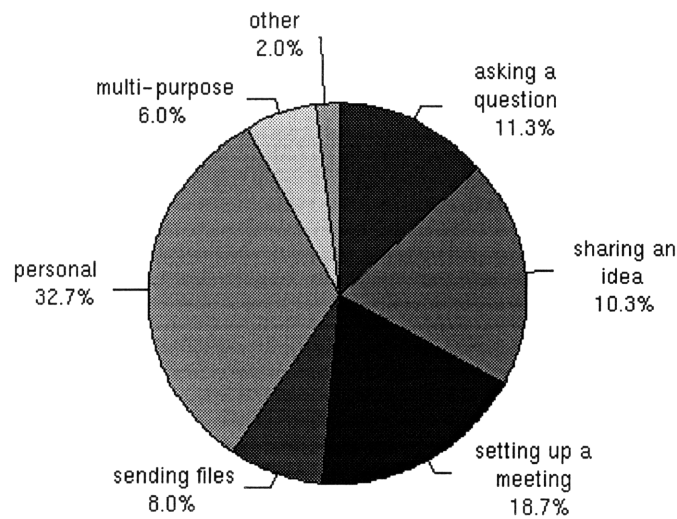


FIGURE 3. Intent of last composed message, students only.



The tests showed that there was a significant difference between students and non-students in the number of personal messages sent. This is illustrated in **FIGURE 2** and **FIGURE 3**. There was also a significant difference between the faculty and the research staff in the number of messages sent with the intent of sharing an idea/opinion. The table below summarizes these results which demonstrate that there are some differences in the intents of messages composed by different groups of people within the Sloan School, a finding which supports Hypothesis 8. Other than those noted below, there were no significant differences in the intents of the messages.

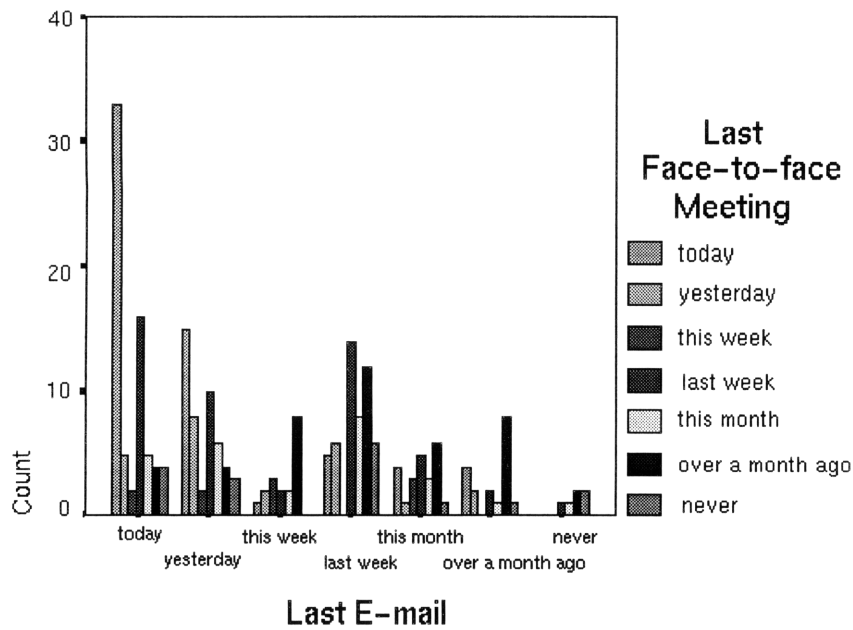
TABLE 4. Message intent: Significant differences

Intent	Position 1	Proportion of messages (1)	Position 2	Proportion of messages (2)	Difference in proportion
Personal	Student	33%	Non-student	6%	$-0.378 < p_1 - p_2 < -0.141$
Sharing an idea/opinion	Faculty	26%	Research staff	0%	$-0.440 < p_1 - p_2 < -0.050$

6.3.4 Hypothesis 1: E-mail v. face-to-face meetings

The last questions in this section asked the respondent to select the last time they sent their most recent correspondent e-mail (excluding the most recent message) and the last time they saw them face-to-face. Their choices for both questions were: today, yesterday, this week, last week, within the last month, over a month ago, and never. I used Spearman's test (2-tailed) for bivariate correlation to see if these two variables were related. The calculated correlation coefficient was 0.349 which is significant at the .01 level. This supports Hypothesis 1 which states that the degree to which people correspond by e-mail depends on how often they see each other face-to-face.

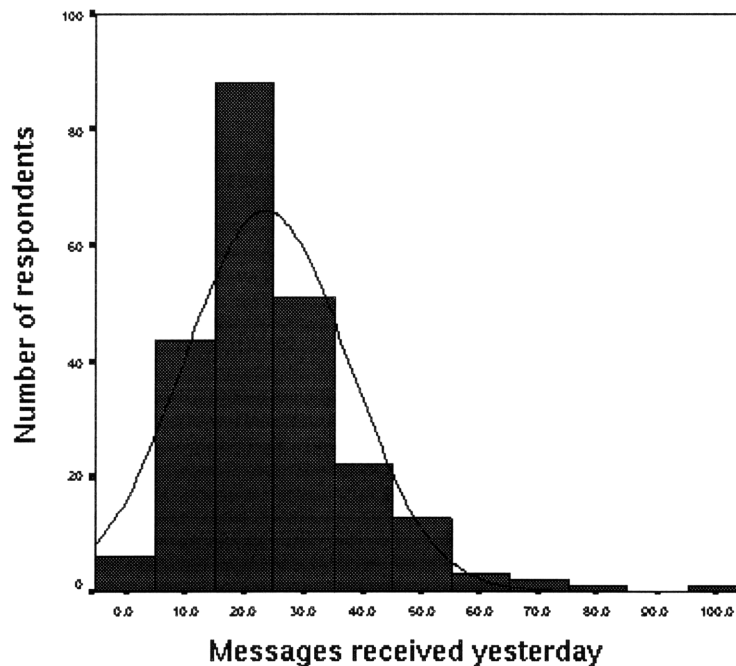
FIGURE 4. Last e-mail v. Last face-to-face meeting



6.4 Messages received within a day

In order to test hypotheses 4 through 7, I collected data on the number of messages the participant received the day before they filled out the survey. **FIGURE 2** shows the histogram for that data (excluding one outlier who received 350 messages) overlaid with the normal curve.

FIGURE 5. Histogram: Messages the respondents received in a day.



The mean was 24.9; the mode and the median were 20. Again, the curve shows that the distribution of messages received is skewed to the right and therefore is not a normal distribution. For this reason, the related hypotheses were tested using Spearman's correlation coefficient (2-tailed).

6.4.1 Analysis of hypotheses 4 through 7

For all of the questions related to the number of messages the participant received during the previous day, the data collected was in ordinal form. The possible choices were: 0-5, 6-10, 11-15, 16-20, and 20+. To analyze the data, those numbers were converted to their median and then those numbers were recalculated as a percentage of the total number of messages received. The interval “20+” was converted to the somewhat arbitrary number 50. Any data point over 100% was omitted from the analysis. **TABLE 5** shows some examples of the data conversion.

TABLE 5. Data conversion examples.

SURVEY RESPONSES		CALCULATIONS		
# of messages received	# of messages deleted	Median of interval	% of messages deleted	Data point included?
32	6-10	8	25	yes
41	20+	“50”	122	no

The catch-all interval of “20+” is some cause for concern. Since we translated this data to have a median of 50, the resulting percentages did not truly represent the proportion of the e-mail that was described by the variable. For this reason, I first analyzed the data including “20+” data points and then excluding them. **TABLE 6** shows the comparison of these analyses which demonstrates that the inclusion of these points actually lends more support to the hypotheses.

6.4.2 Hypothesis 4: Messages read thoroughly

Hypothesis 4 stated that people are less likely to read their messages thoroughly if they receive a large volume of e-mail. For each data point, I calculated the proportion of e-mails read during the previous day to e-mails received. I then analyzed the relationship

between this proportion and the number of messages received using Spearman's correlation coefficient.

The calculated correlation coefficient was -0.420 which is significant at the $.01$ level. This number indicates that there is a significant inverse relationship between the number of messages one receives and the proportion of those messages they read thoroughly-- a statistic which corroborates the hypothesis.

6.4.3 Hypothesis 5: Messages replied to immediately

Hypothesis 5 is that the larger the volume of e-mail one receives, the smaller the proportion of messages one will respond to immediately. The Spearman correlation coefficient for this hypothesis was -0.716 which is significant at the $.01$ level. This shows that there is a strong negative relationship between the variables, an observation which supports the hypothesis.

There is, however, some reason to doubt the correlation of the volume of e-mail one receives to the number of messages one replies to. The survey measured the number of messages the participant replied to, but did not measure how many messages necessitated a reply. The proportion of messages replied to out of messages requiring a reply might have provided a much more insightful statistic.

6.4.4 Hypothesis 6: Deleted messages

Hypothesis 6 states, "The larger the volume of e-mail one receives, the larger the proportion of messages one deletes without reading them." Spearman's correlation coefficient for this hypothesis was calculated to be $-.559$ which is significant at the $.01$ level.

This observation supports the hypothesis by demonstrating a negative relationship between the variables.

6.4.5 Hypothesis 7: Useful messages

Hypothesis 7 is that the larger the volume of messages one receives, the smaller the proportion of message one finds useful to one’s work. The calculated Spearman’s coefficient of -0.449, which is significant at the .01 level, shows a negative correlation between the variables. This supports the original hypothesis.

TABLE 6. Comparison of analyses with and without “20+” data points

Message	Number of “20+” data points	% of sample	Spearman’s <i>r</i> including the data points	Spearman’s <i>r</i> excluding the data points
read	9	3.9%	-0.420	-0.482
replied to	0	0.0%	-0.716	-0.716
deleted	4	1.7%	-0.559	-0.605
useful	2	0.9%	-0.449	-0.473

6.4.6 Junk mail

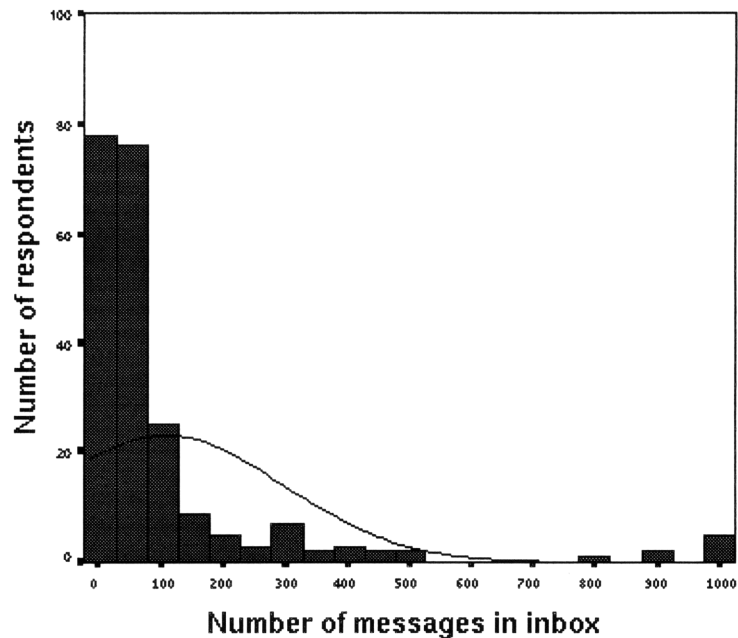
One of the variables measured on the survey was the amount of junk mail the recipient received during the past day. This data was not related to any of the original hypotheses, but it was analyzed with the possibility that it might provide further evidence of the utility of the e-mail the participants received. Again, this data was converted into a proportion of the total number of messages received the day before and then analyzed with Spearman’s bivariate correlation. The resulting correlation coefficient was 0.578 which is significant at the 0.01 level. This implies that the larger the volume of mail received, the greater the amount that is considered junk mail.

6.5 Inbox composition

Earlier in my research design, I made the assumption that the messages people kept in their inbox were the ones they felt most required their attention. Most of the people pre-surveyed deleted those messages that were not important or filed away messages that they might need later in other folders. The messages they kept in their inbox were the ones they wanted to be sure to look at when they checked their mail. By looking at the number of messages a respondents had in his/her inbox, I hoped to see whether or not people were indeed keeping a manageable volume of e-mail.

FIGURE 1 is a histogram of the number of messages the respondents had in their inboxes (excluding the 8 outliers who had over 1,000 messages). Overlaid on the histogram is the related normal curve. As the curve illustrates, the data is heavily skewed to the right-- not at all a normal distribution.

FIGURE 6. Histogram: Total number of messages in inbox.



I was somewhat surprised that the results covered such a wide range. The total number varied from 0 to 7800 messages, with the median being 40 and the mode being 50. The fact that 63 of the respondents (27%) kept more than 100 messages in their inboxes casts some doubt on my original assumption that people keep important messages they want to pay attention to in their inboxes since that is not a very manageable volume.

The faculty in particular kept a large volume of e-mail in their inboxes. Their average was 427 messages; the research staff's was 15; the support staff, 65; the students', 147; and the administrative staff, 222.

One of the respondents with a large amount of e-mail in his inbox commented, "I tend to be sloppy about creating folders to save mail I may need later, and instead let it accumulate in my inbox. It is rare that I go back and read them, and monthly I tend to just throw most of them out." This e-mail management approach may explain why so many of the respondents retain a lot of messages.

7.0 Results

The analysis of the data can be summarized with the following results:

- People at the Sloan School receive an average of 25 messages each day.
- Students send significantly more personal messages than any other group. This is to be expected since in a professional environment, employees are discouraged from spending their time on personal issues whereas students have no such restrictions.
- The volume of e-mail one receives does significantly affect how much attention is paid to each message. As the volume of e-mail increases, people delete a higher proportion of messages without reading them, read a smaller percentage all the way through, and find a smaller amount to contain useful information. Also, the more mail people get, the larger the percentage of it they think is junk.
- There is a significant positive correlation between the degree to which people communicate with e-mail and the frequency with which they see each other face-to-face.
- People at the Sloan School keep such a varied amount of messages in their inboxes that the original assumption that people keep messages they want to pay close attention to was rejected. Twenty-seven percent had over 100 messages in their inbox, a number which casts some doubt on how carefully those messages are being managed. Even so, the data was analyzed to see if people kept certain types of messages in their inboxes. The results showed that no type of message (report, meeting reminder, technical description, personally addressed message, intra-departmental message) was retained significantly more than others.

8.0 Commentary

8.1 Problems encountered

8.1.1 Ambiguity

It is ironic that most of the problems encountered during this research project resulted from ambiguity-- one of the problems identified in the research as a weakness of electronic communication. The e-mail soliciting responses to the survey was sent out to the Sloan School mailing list on a Monday. Since three of the questions referred to events that occurred "yesterday" some of the respondents were confused as to whether or not they should respond about their e-mail activities on Sunday or on Friday. Clearly, the volume of work- or school-related e-mail is significantly less on a Sunday than on a weekday. To rectify this, the data from those who stated that they answered the questions about Sunday were omitted from the analysis.

Another question asked the participants to enter the amount of "junk mail" they received. Again, the respondents had a few different ideas of what junk mail is. Some felt that any mail not personally addressed to them was junk. Others described it as e-mail from people they did not know or any unsolicited e-mail. Other things classified as junk mail were: jokes, postings to the Sloan School mailing list, advertisements, and repetitive e-mails.

8.1.2 The limitations of multiple-choice

The weakness of this sampling plan was that some of the multiple-choice ques-

tions attempted to encompass a broad range of possible responses with a few choices. For example, the question that asked the intent of the last e-mail the respondent wrote lists as the choices, “asking a question, setting up a meeting, sending files, sharing an idea/opinion, personal, multi-purpose, other.” Clearly these selections do not represent all the possibilities. An alternative to the multiple-choice questions would have been to ask open-ended questions (of the type: “Fill in below the intention of your last e-mail.”). However, I felt this type of question would reduce the response rate and the results yielded would be difficult to quantify and largely inconclusive.

8.1.3 Interval questions

One of the survey questions asks the respondents the number of messages they deleted. During the pre-survey, it was determined that it was hard to remember the exact number, so instead of leaving the question open-ended, it was formatted as a multiple choice between different intervals. To illustrate:

When you first read your e-mail today, approximately what number of those messages did you delete without reading?

- 0-5 6-10 11-15 16-20 20+

The problem with this format is that it is only appropriately descriptive when most of the respondents received more than 20 messages and deleted less than 20. For those who received (for example) 200 messages and deleted 100 or more, their response does not shed much light on the proportion of e-mail that they felt was unnecessary. Similarly, for those people who only received 5 messages and deleted 1 of them, the resulting data

point is not very useful. To account for this, I analyzed the interval questions with non-parametric statistical tests that require only ordinal measurements.

8.1.4 Differences in e-mail software

One of the assumptions about the sample group was that they used e-mail software with similar features. This proved to be an incorrect assumption. Some e-mail software offers features such as inbox filtering which automatically distributes messages that fit certain specifications into different folders. People who use this feature relinquish their control over which messages they keep in their inbox, so all of the questions that asked about the messages in the participant's inbox became ambiguous. Two of the respondents cited this difference.

8.2 Possible improvements

Most of the problems cited above could probably have been avoided with a more thorough pre-survey. Since the pre-survey was directed at a limited number of people (all of whom were students), it did not catch many of the ambiguities that became apparent when the sample group responded to the survey.

Another improvement to this survey would have been to ask more open-ended questions about the respondents' use and perception of e-mail. Data of this sort might have helped in the identification of special cases and outliers, among other things.

8.3 Recommendations

After conducting this research study, several ideas came to mind about possibilities for further study.

1. It would be interesting to determine how e-mail is used relative to other modes of communication. What types of issues do people choose to discuss over the phone or face-to-face via e-mail?
2. It was determined that the degree to which people communicate by e-mail is related to how often they see each other face-to-face. Why is this so?
3. Is there any method of e-mail management that maximizes the amount of information gained per e-mail read?
4. It was noted that the faculty keep significantly more messages in their inbox (on average, 426) than other groups within the Sloan School even though they receive just about the same number of messages each day (on average). Why is this so?
5. Many of the respondents commented that they felt the Sloan School mailing list was abused and added greatly to the amount of junk mail they received. This alludes to the problems that currently exist with the management of large mailing lists and the types and contents of messages sent to them. A recommendation would be to develop a protocol to regulate these mailing lists to see whether or not these problems can be minimized.

9.0 Conclusion

The intent of this research project was to study how people use e-mail in an academic setting. As a result of this study, I have determined some of the ways people manage their e-mail and some of the differences particular to groups of people within the Sloan School. Seven of the eight original hypotheses were supported by data from the survey, but further study is needed to explore these results and determine why they are so. (It may be that the sample is biased in that the types of people who would respond to the initial e-mail request to complete the survey have a certain common mentality which was evidenced in the results.)

E-mail is a useful, efficient means of communication that has changed how, what, and with whom we communicate. Many of the survey respondents commented that their jobs would be impossible without it. It is important, however, that we continue to pay attention to the ramifications of its use, overuse, and misuse so that e-mail continues to be an effective means of communication.

10.0 Bibliography

- Allredge, Richard Warren. *Electronic message systems: Factors affecting their acceptance*. MIT Thesis, 1980.
- Anderson, Robert H. et al. *The design of the MH Mail System*. Santa Monica, CA: Rand, 1989.
- Anderson, Robert H. et al. *Universal access to e-mail: feasibility and societal implications*. Santa Monica, CA: Rand, 1995.
- Bellinger, William Hagood. *Electronic mail systems: are they effective in the office?* MIT Thesis, 1980.
- Blum, Daniel. *The e-mail frontier*. Reading, MA: Addison-Wesley Publishing Company, 1995.
- Comins, Christine. *Effects of an electronic message system on office communication patterns*. MIT Thesis, 1979.
- Connell, Stephen. *Electronic mail: a revolution in business communications*. White Plains, NY: Knowledge Industry Publications, 1982.
- Cummings, William Richard. *Strategic planning in the public sector: the case of electronic mail and the postal service*. MIT Thesis, 1980.
- Dixon, Wilfrid J. and Frank J. Massey, Jr. *Introduction to Statistical Analysis*. New York, NY: McGraw-Hill Book Company, Inc., 1957.
- Shapiro, Norman and Robert H. Anderson. *Towards an ethics and etiquette for electronic mail*. Santa Monica, CA: Rand, 1985.

Sproull, Lee. *Connections: new ways of working in the organization*. Cambridge, MA: MIT Press, 1991.

SPSS Base 7.0 Applications Guide, SPSS Inc., 1996.

SPSS Base 7.0 for Windows User's Guide, SPSS Inc., 1996.

Star, Susan Leigh. *The cultures of computing*. Oxford, UK: Blackwell Publisher, 1995.

Vervest, Peter. *Innovation in electronic mail*. New York, NY: North Science Publishing Company, 1987.

11.0 Appendices

11.1 The Survey

FIGURE 7. Screen capture of the survey, part 1 of 3.

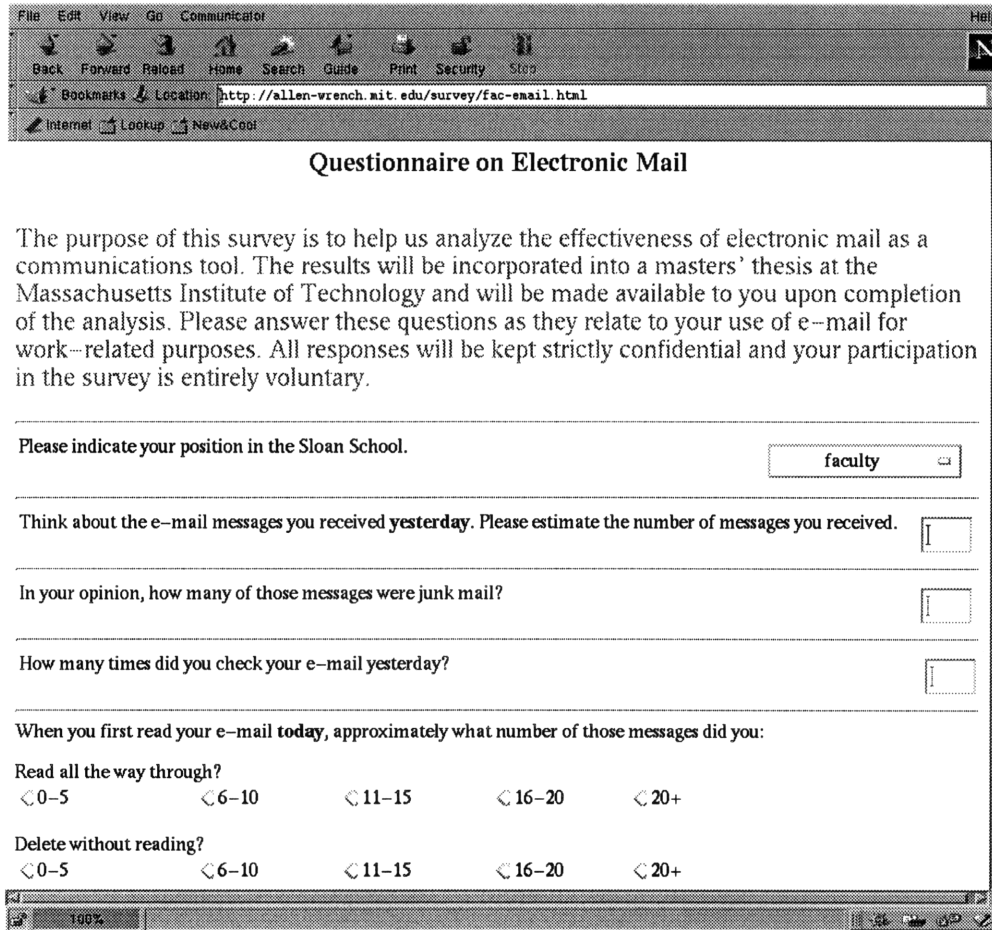


FIGURE 8. Screen capture of the survey, part 2 of 3.

File Edit View Go Communicator Help

Back Forward Reload Home Search Guide Print Security Stop

Bookmarks Location: <http://allen-wrench.mit.edu/survey/fac-email.html>

Internet Lookup New&Cool

Respond to immediately?

0-5 6-10 11-15 16-20 20+

Find useful to your work in some way?

0-5 6-10 11-15 16-20 20+

How many messages are currently in your inbox?

Of the messages currently in your inbox that you received today, how many fall into the following classifications:

A meeting reminder/request?

A technical description of something which may be useful to you?

A report?

A message from someone within your department?

A message addressed to you personally (rather than to a mailing list)?

Please think about the last person to whom you **initiated** an e-mail message (excluding messages addressed to more than one person or responses to messages that you received). For all the questions below, please select the answer that best fits.

Where is the recipient located?

Standard list Standard stop com FrameMaker Jds/shona WorkMan NistScene zenbur

FIGURE 9. Screen capture of the survey, part 3 of 3.

File Edit View Go Communicator Help

Back Forward Reload Home Search Guide Print Security Stop

Bookmarks Location <http://allen-wrench.mit.edu/survey/fac-email.html>

Internet Lockup New&Cool

What is his/her affiliation to you?

What was the intent of the e-mail you sent?

Now think about the last time you sent or received an e-mail message from this person (excluding the message referred to in the question above). How long ago was that?

When was the last time you spoke to this person face-to-face?

Please add any comments here.

Thank you for taking the time to complete this survey. Press to submit your answers.

100%

11.2 Data Tables

11.2.1 Hypothesis 1

TABLE 7. Last e-mail v. last face-to-face meeting

Last face-to-face meeting								
Last E-mail	Today	Yester-day	This week	Last week	Within the last month	Over a month ago	Never	TOTAL
Today	33	5	2	16	5	4	4	69
Yesterday	15	8	2	10	6	4	3	48
This week	1	2	3	2	2	8	0	18
Last week	5	6	0	14	8	12	6	51
Within the last month	4	1	3	5	3	6	1	23
Over a month ago	4	2	0	2	1	8	1	18
Never	0	0	0	1	1	2	2	6
TOTAL	62	24	10	50	26	44	17	233

Spearman's rho correlation coefficient: 0.349, significant at the .01 level (2-tailed).

11.2.2 Hypothesis 2

TABLE 8. Frequencies: Location of correspondent

	Adminis-trative staff	Faculty	Research staff	Student	Support staff	TOTAL
Same floor	7	4	0	11	4	26
Same building	5	3	0	26	1	35
Different building	10	6	4	37	8	65
Outside MIT	5	18	2	76	6	107
TOTAL	27	31	6	150	19	233

TABLE 9. Proportions: Location of correspondent

	Adminis- trative staff	Faculty	Research staff	Student	Support staff	Non- student
Same floor	0.259	0.129	0.000	0.073	0.211	0.183
Same building	0.185	0.097	0.000	0.173	0.053	0.110
Different building	0.370	0.193	0.667	0.247	0.421	0.341
Outside MIT	0.185	0.581	0.333	0.507	0.316	0.378

11.2.3 Hypothesis 3

TABLE 10. Frequencies: Affiliation of correspondent

	Adminis- trative staff	Faculty	Research staff	Student	Support staff	TOTAL
In group	97	6	1	32	5	53
In area	1	0	1	2	2	6
Within Sloan	7	4	0	48	4	63
Within MIT	6	4	2	13	2	27
Outside MIT	4	16	2	55	6	83
TOTAL	27	30	6	150	19	232

TABLE 11. Proportions: Affiliation of correspondent

	Adminis- trative staff	Faculty	Research staff	Student	Support staff	Non- student
In group	0.333	0.200	0.167	0.213	0.26	0.256
In area	0.037	0.000	0.167	0.013	0.105	0.049
Within Sloan	0.259	0.133	0.000	0.320	0.211	0.183
Within MIT	0.222	0.133	0.333	0.087	0.105	0.171
Outside MIT	0.148	0.533	0.333	0.367	0.316	0.341

11.2.4 Hypotheses 4 - 7

TABLE 12. Management of messages received yesterday (Spearman's *r*)

		# Received yesterday	Deleted	Read through	Replied to	Useful
	# Received yester day	1.000	-.559*	-.420*	-.716*	-.449*
PROPORTION	Deleted	-.559*	1.000	.149	.338*	.149*
	Read through	-.420*	.149	1.000	.558*	.434*
	Replied to	-.716*	.338*	.558*	1.000	.525*
	Useful	-.449*	.149	.434*	.525*	1.000

* Correlation is significant at the .01 level (2-tailed).

11.2.5 Hypothesis 8

TABLE 13. Types of messages in inbox (Spearman's *r*)

Type of message	Inbox total
Intra-departmental	.017
Meeting reminder	-.047
Report	-.001
Technical description	-.013
To person (not group)	.034

* None of these coefficients is significant at the .01 level (2-tailed).

11.3 Comments from the respondents

- E-mail is a valuable tool which we use to keep in constant communication with students; participants in international projects; and keeping in touch with our MIT colleagues.
- The last message I sent was somewhat unusual in that I was reporting a scam to Purchasing - not thank heavens an every day event! As far as I am concerned e-mail is the greatest thing since sliced bread! Saves me a lot of time and aggravation and motivates me to do things I might not get around to if they required actually getting someone on the phone or seeing them - like reporting a scam!
- E-mail is an integral part of my job. I use it to communicate both with colleagues whom I see every day and participants for courses that we offer writing from across the globe. I find it an essential tool used to stay in touch with my boss when she travels. One of the best technological advancements our office has made lately was to add an on-line application form for our Summer Session course to our web site. The applications are received in our office via e-mail and we reply immediately to the applicant via this inexpensive and convenient mode of communication. Many of them send thank-you notes and praises for the prompt reply and the high level of customer service that we offer all made possible by e-mail.
- When I am in a hurry I scan the messages in my inbox for personal stuff and leave the others for later. Some messages hang around for months because they contain an address etc. that I want to keep.
- I would be unable to do my job effectively without e-mail.
- I have always preferred e-mail over [voice-] mail. I've been using e-mail for over 30 years, and] I still find it an effective and convenient communications tool.
- I respond/file/dispose of all personal messages right away and leave others in there for reminders.
- Very interesting and thought provoking. I never use filters with my e-mail because I still do not trust in the technology and I am afraid to lose something unintentionally. E-mail has been a powerful tool for both of us (we're doing a joint thesis) up to the point in which we only have to meet once a week.
- I receive numerous e-mail inquiries for a particular program (anywhere from 50-350 per week), requesting info about the program or other info. I can barely keep up with the volume; e-mail requests have cut down on phone inquiries but added a whole other level of work. Amidst all of that, I probably receive about 10-30 other messages per week. I often dread opening my e-mail, especially during our busiest times of the year.
- I find with the overkill on e-mail, the personal approach, for example phone call, bumping into someone in the hallway etc. is gaining it's effectiveness...
- E-mail is very useful for my work... I use it a lot to keep up to date with friends back home and those around me.

- I can't live without e-mail...
- I didn't think I had a problem until my e-mail listing was added to the Sloan students' list. Now most of my e-mail is random mass mailings that waste my time to sort through. I am on many e-mail lists, but none is so abused as the Sloan list.
- I think there is an opportunity for some start-up company to develop software to recognize junk mail and give users the option not to download unwanted e-mail.
- E-mail lacks "nuance" of expression. It's easy to upset folks unintentionally.
- I tend to be sloppy about creating folders to save mail I may need later, and instead let it accumulate in my inbox. It is rare that I go back and read them, and monthly I tend to just throw most of them out.
- You have to take the good with the bad. E-mail is vital to today's virtual and fast-paced working environment. It actually doesn't take too long to weed through junk mail once you are accustomed to looking for what you want. At work, I usually get between 75-100 e-mails per day, most of which are junk e-mails. But the ones that are work related are vital to business communications, particularly the ones that are off-site.
- Being a student, I mostly use e-mail for personal use with friends and family, in addition I am using it to handle the details of my new job assignment as well as to be notified of events around school.
- Although I may classify many e-mails as "junk" mail, many times they still contain information that I may be interested in. For example, I may subscribe to a list of some sort that 90% sends out info that I'm not interested in. I'm willing to wade through the junk however because 10% of the time there is something of interest to me.
- Most of the e-mails I receive are work related - faculty and staff looking for information and results. I am a video conference project coordinator. My e-mail address and requests are now global. Very challenging!
- I don't ever delete a message without first reading it through.
- I find that the best way I can keep my inbox free of clutter is to read and discard messages the same day. I usually keep only those messages I deem important, the rest are deleted. Additionally I move e-mails that are older than 1 month to an "old mail" folder. I check my e-mail more frequently than my voice message service.
- It's an essential communication tool, no doubt about it, but users need to be conscious and not abuse the system.
- I use e-mail for a variety of qualitatively different purposes -- getting news, keeping in touch with family, working on dissertation, communicating with professors with whom I do research or TA, etc. Within each category, the percentage of useful e-mails is VERY different. Lots of the junk mail is actually from people in Sloan sending e-mail to the mailing lists. The rest are mostly jokes that are sent by friends.
- I love e-mail. It's such a fantastic, efficient way to communicate. I like it more than voice-mail because you don't have to worry about "sending unintentional messages" that someone reads from your tone of voice. I'm addicted, in particular because it is so

useful. Before I was a student at Sloan, we had it at work, but I used it less and it just didn't seem essential. At Sloan, it is essential. I don't mind nuisance mail all that much. It's very easy to delete.

- I never delete an e-mail without reading the first line or so along with the header and the name of the sender. A lot of e-mail I don't read all the way through however.
- Lots of the faculty I work with are often on the road or working in their home office. E-mail is invaluable for quick info exchanges with them.
- A good way to improve the usefulness of e-mail is to provide a longer descriptive header for easier filtering/sorting/deleting. Also, the excess use of schoolwide distribution lists for items that are of only interest to a handful of people, adds to the 'overflowing inbox' syndrome. It becomes easier to just delete large blocks of e-mail vs. reading each message individually for useful information.
- There is way too much e-mail usage going on at Sloan that is non-value added e-mail. I have accidentally deleted e-mail addressed to me personally because I thought it was junk, and there is no way to tell if it is addressed just to you personally. I think e-mail can be a very valuable tool, but more often than not these days, it has become a burden--especially when people send out 3-4 reminders for the same event.