1

National Library
of Canada

Bibliothèque nationale
du Canada

CANADIAN THESES
ON MICROFICHE

THÈSES CANADIENNES
SUR MICROFICHE

0-315-04875-1

52769

NAME OF AUTHOR/NOM DE L'AUTEUR John Ellis

TITLE OF THESIS/TITRE DE LA THÈSE The Effects of System Response Time
on a Word Processing Task

UNIVERSITY/UNIVERSITÉ Carleton University

DEGREE FOR WHICH THESIS WAS PRESENTED/
GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE M.A.

YEAR THIS DEGREE CONFERRED/ANNÉE D'OBTENTION DE CE GRADE 1981

NAME OF SUPERVISOR/NOM DU DIRECTEUR DE THÈSE Dr. Richard Dillon

Permission is hereby granted to the NATIONAL LIBRARY OF
CANADA to microfilm this thesis and to lend or sell copies
of the film.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈ-
QUE NATIONALE DU CANADA de microfilmer cette thèse et
de prêter ou de vendre des exemplaires du film.

The author reserves other publication rights, and neither the
thesis nor extensive extracts from it may be printed or other-
wise reproduced without the author's written permission.

L'auteur se réserve les autres droits de publication; ni la
thèse ni de longs extraits de celle-ci ne doivent être imprimés
ou autrement reproduits sans l'autorisation écrite de l'auteur.

DATED/DATE

Jan 16/81

SIGNED/SIGNÉ

J. Ellis

PERMANENT ADDRESS/RÉSIDENCE FIXE

12 Melgund Ave Apt: 2

Ottawa Ontario

Canada

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QU'ELLE
NOUS L'AVONS REÇUE

EFFECTS OF SYSTEM
RESPONSE TIME
ON A WORD PROCESSING TASK

by

C

John K. Ellis, B.A. (Honours)

A thesis submitted to the Faculty of
Graduate Studies in partial fulfilment
of the requirements for the degree of
Master of Arts

Department of Psychology
Carleton University
Ottawa, Ontario
January, 1981

The undersigned recommend to the
Faculty of Graduate Studies
acceptance of the thesis
"Effects of System Response Time
on a Word Processing Task"

submitted by
John K. Ellis, B.A. (Honours)
in partial fulfilment of the
requirements for the degree of
Master of Arts

.....
Thesis Supervisor

.....
Chairman, Department Psychology

Carleton University

January, 1981

ABSTRACT

Positively skewed distributions were generated by a Bernoulli process with two parameters, the probability of service and minimum delay to provide system delays on a simplified word processing task. Subjects were required to correct single lines of text presented on a CRT using a cursor control box and keyboard. The purpose of the study was to relate these two parameters to user satisfaction with the system, and user performance on the system. It was found that the parameter minimum delay was related to user satisfaction, with rated satisfaction decreasing with increases in minimum delay. Minimum delay was also found to have an effect on user performance, with some evidence that subjects actively compensate for system inadequacies by using different task strategies. Probability of service had similar effects on subjects' satisfaction and performance, although these effects were not as strong as for minimum delay. There was no evidence of a probability X delay interaction. Also, the "Boies effect", the pacing of user work speed by system response time was not replicated. Suggestions were made for possible changes in the experimental task to make it more psychologically demanding. It was reasoned that such changes might provide better conditions for the Boies effect to occur.

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Dr. Richard Dillon, for his assistance and guidance throughout this project. A special thanks goes to Dr. Jo Tombaugh. That this project could have been completed without her help is simply unimaginable. Finally, a very special thanks to Julie for her understanding and patience.

TABLE OF CONTENTS

	Page
INTRODUCTION.....	1
MATERIALS AND METHOD.....	13
Subjects.....	13
Apparatus.....	15
Experimental Task.....	16
Practice Sentences.....	17
Experimental Sentences.....	17
Instructions.....	18
Design.....	19
Dependent Measures.....	19
Other Measured Variables.....	20
RESULTS.....	22
Effects of Probability of Service.....	26
Effects of Minimum Delay.....	29
Interaction Between Probability of Service and Minimum Delay.....	32
DISCUSSION.....	33
REFERENCE NOTES.....	38
REFERENCES.....	39

APPENDICES •

Page

A	General Information Questionnaire.....	41
B	Practice Sentences.....	42
C	Experimental Sentences.....	43
D	Experimental Instructions.....	48
E	Satisfaction Rating Scale.....	53
F	Outliers.....	54
G	Number of Errors Made by Subjects.....	58
H	Univariate Analyses.....	61
I	Tables of Discriminant Weights.....	74
J	ANOVA Summary Tables.....	82
K	Trend Analyses Summary Tables.....	91
L	Raw Data.....	97

LIST OF TABLES IN TEXT

Table		Page
1	Expected Means and Standard Deviations on System Response Time.	11
2	Obtained Means and Standard Deviations on System Response Time.	23
3	Number of Subjects With No Errors for the Experimental Conditions.	25
4	Structure Coefficients for First Discriminant and Univariate Fs for the Probability of Service.	28
5	Structure Coefficients for First Discriminant for the Minimum Delay.	31

LIST OF FIGURES IN TEXT

Figure

Page

1

Representation of the model
for generating system
response times.

9

INTRODUCTION

When a user enters information into a computing system and anticipates a reply, there may be a noticeable delay, called the system response time, before the system responds. Over the duration of a user's computer session a distribution of delays will take shape. Interest in this distribution stems from the belief that certain of its characteristics can influence user performance and satisfaction. For example, Doherty (1979) has claimed that a mean delay of only two seconds would result in a loss of 10,000 man-hours per month for his associated laboratory. This impressive estimate was based on the so-called "Boies phenomenon"; the slowing of work speed with increases in mean delay length, (the Boies phenomenon will be discussed later).

Given the psychological and economic implications of system response time, one would expect a great deal of empirical research devoted to the topic. However, a recent and extensive review of the human factors, and human/computer interface literature by Ramsey and Atwood (1980), revealed only five studies on the general topic of system response time. Since the importance of system response time was being espoused as early as 1968 by several authors (Miller 1968, Carbonell, Elkind & Nickerson, 1968), a greater accumulation of research findings might have been expected. Instead, there is an absence of basic facts, and unanswered questions.

While there has been an apparent lack of research, there have been past efforts at providing guidelines for desirable system response time characteristics. Two of these will be considered: the contribution of Miller (1968) and the interesting paper by Carbonell et al. Both papers provide insight into which aspects of system response time are considered important. Because Miller's paper has had greater impact on the computer industry, it will be considered first.

Miller, using a psychological model, developed a set of maximum acceptable mean delays for various categories of computer commands. Viewing human-computer interaction as a problem solving situation the model emphasizes the need to minimize disruption of ongoing problem solving processes. The model primarily involves three concepts: (1) short term memory, (2) spontaneous problem organization, and (3) closure.

Short-term memory, the foundation of the model, is characterized as a temporary memory store for limited amounts of information. Borrowing Klatz's (1977) analogy, the model's version of short-term memory can be understood as a workbench where tools and materials needed for current tasks are kept; similarly in human/computer interaction, information required for problem solving activities must first be stored in the user's short-term memory, and subsequently accessed from short-term memory to be used.

The model assumes that individuals spontaneously organize a problem for solution into what Miller called clumps, or in more current terminology sub-tasks, or sub-problems. The problem organization is itself assumed to be a hierarchical sequence of sub-tasks, with progression through the sequence bringing the problem closer to, and eventually to, solution. As an example of problem solution organization, consider the sequence of sub-tasks based on Miller's example of "making a telephone call":

1. find the number in the telephone book
2. remember the number
3. dial the number
4. say "hello"

The completion of each sub-task results in the individual experiencing "a sense of temporary completion" (Miller, 1968, p. 268) called closure, with the amount of closure experienced being an increasing function of sub-task depth in the hierarchy. In the example above, closure after "remembering the number" would be greater than after "finding the number in the telephone book", but less than after "dialing the number". Within the model, closure is given the status of a human need.

Because information necessary for sub-task solution is stored in short term memory, forgetting from this memory can have obvious negative consequences. According to the model, the probability of forgetting increases as the interval between the start of a sub-task and closure increases, or in other words, as some function of time. The probability increases further as an individual begins to experience an awareness of waiting, a result of interference from accompanying thoughts, which presumably begin to occur within about two seconds. If forgetting occurs after closure, it has negligible consequences since closure implies sub-task completion, and in fact, the model assumes that short term memory is at least partially cleared when closure occurs. Since closure is a need, delaying its gratification results in frustration. However, tolerance for delays increases with the amount of closure anticipated to follow a delay.

Miller used his model essentially as a heuristic to suggest maximum allowable system response times for different categories of computer commands. He hypothesized that the relationship between performance, and the maximum tolerable delay for a given activity is a step function; once the tolerable delay has been exceeded a dramatic effect on performance is expected. This performance decrement is generally attributed to some sort of disruption of a problem solving

process, e.g., forgetting. While Miller did not specify a maximum tolerable variance for system response time he did suggest a "range rule". The rule, applicable to mean delays falling within two to four seconds, states that the range of system response times should not exceed $\pm$ eight percent of the mean delay. This rule was derived by Miller from his evaluation of psychophysical studies (Stevens, 1951) on the perception of stimulus durations, which showed that variable durations within the intervals specified could not be discriminated. According to Miller's reasoning, delays obeying the range rule are perceived as the same, and therefore psychologically do not exceed the maximum tolerable delay.

Miller's model, and therefore his recommendations are based on his understanding of then current (1968) psychological concepts, plus the extrapolation of some laboratory results. The approach does have intuitive appeal (from a psychological perspective) and some face validity; similarities are apparent between some experimental paradigms and computer tasks, e.g., some aspects of debugging a program and letter search tasks. However, while of heuristic value neither the model nor its assumptions have been seriously assessed empirically. In addition, the model has not been revised since first formulated by Miller (1968) to reflect the impressive

amount of research done to date on memory and problem solving processes. Also, as will be discussed, the mean and the variance are not best suited to characterize the system delays. Consequently, Miller's model will not be directly tested in this experiment.

Carbonell, Elkind & Nickerson (1968) advocate an empirical modelling approach to study the "importance of time in a time sharing system", and although they do not offer a specific model themselves, they demonstrate the approach using hypothetical examples. Their examples provide conjectures as to relevant dependent and independent variables. Included among the independent variables are system response time characteristics. In general, Carbonell et al. suggest that shorter mean delays are preferable, with the maximum tolerable delay for simple commands assumed to be within milliseconds. The maximum tolerable delay is thought by these authors to increase with command complexity. Presumably, expectations developed by users are related to a judgment of the amount of time required by the computer to complete various commands. However, users are also thought to prefer longer mean delays to shorter mean delays when the former are less variable than the latter, because they are more predictable by the user.

Two criterion measures are suggested to assess the effects of system response time on a computer user: (1) the

value of work accomplished on line, relative to the value of work achieved otherwise, and (2) user ratings of satisfaction with time related aspects of a system. Essentially, Carbonell et al. suggest determining the relationship between the independent and dependent variables, and developing a response surface model. Their discussion is largely "exploratory", and their recommendations are largely based on personal observations and experience.

The recommendations of Miller and Carbonell et al. share an emphasis on the importance of short mean system delays, and the importance of the consistency of delays.

Empirical studies done tend to corroborate Miller's and Carbonell et al.'s contention that system response time is a psychologically important system attribute. Boies (1974) studied the effects of system response time on users interacting with a computer via teletype terminals. A substantial positive correlation ($r = .84$) was found for system response time and the response time of users. This is the so-called Boies phenomenon - the apparent pacing of user work speed by system response time. Abrahamson and Murray (1980) found that system delays generated from a rectangular distribution lowered the rate of user interaction on a simulated video text system (i.e., a Telidon like system). They found that the delays increased user errors and that a variable delay was worse than a constant delay.

The work reviewed has emphasised manipulation of the mean and variability of delays as the approach to studying the effects of system response to time on user performance and satisfaction. The problem with this approach is that the mean and variance do not easily relate to the processes that generate the delays. Thus it is not easy to relate these parameters to actual systems. To avoid this major limitation, system delays were generated for this experiment, in the context of a word processing task, by a model with parameters corresponding to actual timesharing systems. The model used to generate the system delay following each response is shown in Figure 1. After a minimum delay of length M , a decision to service or postpone service to the user is made by independent sampling from a Bernoulli distribution in which the probability of service (p) is stationary over trials. Thus, the user is serviced after a delay of M with probability p , or after a delay of $2M$ with probability $p(1-p)$ etc. For example, when a subject presses a cursor key the minimum delay (e.g., .2 sec.) occurs. Next, a decision is made to execute or not execute the user's cursor command. If the decision is made not to execute the command, then another minimum delay occurs so that the system delay is now 2 times the minimum delay or .4 seconds. This process continues until a decision is made to execute the command,

Figure 1. Representation of the model for generating system response times.

this decision being followed by the response of the system. The expected mean delay is M/p , while the expected standard deviation is $M/p\sqrt{1-p}$. The values of M , p and the expected mean delays and standard deviations for the conditions used in the experiment are shown in Table 1.

There were two major reasons for selecting this model, to vary delays. First, the theoretical distributions generated are similar in form to those observed when measuring the response time of Telidon in a natural setting (Tombaugh & Dillon, 1980). That is, system response times are represented as a positively skewed distribution with the minimum pause equal to the modal pause. Thus, a very simple sampling model approximates complex conditions.

Second, information about the relationship between minimum system delay, probability of service, and user responses was judged to be of greater potential utility to system designers and managers than information based directly on mean delays and standard deviations. Mean delays and standard deviations do not independently correspond to particular system characteristics in a typical time sharing situation. That is, a change in the system will probably affect both the mean and standard deviation of the response time. On the other hand, the minimum delay and the probability of service can be meaningfully linked to specific attributes of the system. For example,

Table 1

Expected Means and Standard Deviations
On System Response Time

<u>Condition</u>			
Probability of Service	Minimum Delay	Mean	Standard Deviation
.25	.2	.8	.69
	.4	1.6	1.39
	.8	3.2	2.78
.50	.2	.4	.28
	.4	.8	.57
	.8	1.6	1.14
1	.2	.2	0
	.4	.4	0
	.8	.8	0

questions concerning the number of users allowed access to a system can be related to the minimum delay. Multiplying the maximum number of users by the time slice dedicated to each user gives a rough estimate of minimum delay under worst case conditions. In contrast, the probability of service relates to the probability of avoiding system bottlenecks, such as waiting for disk access, or to being postponed for priority scheduling. Thus, the addition of another printer to reduce queuing would increase the probability of service.

While the model is highly simplified, it is a heuristic device which a system designer or manager can use to link decisions about system characteristics to user reactions. The utility of this approach depends upon (1) establishing links between the model parameters and the behavior of users, which is the purpose of this research and (2) determining the degree to which changes in system characteristics affect the system response times in a manner compatible with the model. The experimental task was a simplified word processing task in which subjects were required to correct single lines of text. The task has face validity and is realistic enough to provide some basis for the generalization of results. As well, word processing is a growing computer application so that the study will be of interest to both users and designers of word processing systems.

MATERIALS AND METHOD

Subjects. There were 90 subjects, randomly assigned in equal numbers to the nine experimental conditions. The data for five subjects were lost due to equipment failure, while one subject's data was dropped from the analysis because of his apparent difficulty understanding the rating scale instructions. A second subject's data was eliminated because he revealed to the experimenter, in a post-experimental interview, that at some point during the experimental session he had left the computer terminal to look out a window. These subjects were replaced by seven additional subjects.

The subjects were recruited from several sources: Canada Student Manpower (57 subjects), introductory psychology class (7), respondents to posters placed in various university locations (5), respondents to an ad in a local newspaper (10), psychology graduate students and non-academic employees of the psychology department (11).

All subjects were paid five dollars for their participation, with introductory psychology students also getting course credit.

The subjects were 41 males and 49 females with a combined mean age of approximately 21 years. The approximate modal and median ages were 18 and 19 yrs., respectively, with

the range from 15 to 65 years. The age statistics are based on 89 subjects; one subject choosing not to reveal her age.

To assess the extent of subjects' prior computer experience, subjects were asked to provide information concerning their program major, number of computing science courses, and the number of hours spent interacting with a computer. However, for subjects from the primary source, Canada Student Manpower, the first of these questions was inappropriate since the majority of these subjects were not university students. The question was unanswerable for a majority of these subjects so is not considered further.

Most subjects, 77.8 percent, did not have a single computing course to their credit, while 12.2 percent had one such course, 7.8 percent had two courses, and two subjects had three and five courses, respectively. Subjects were asked to estimate the number of hours they had spent interacting with a computer. The majority of subjects, 65.6 percent, had one or fewer hours interaction. At least five subjects (5.6 percent), had at one time, or were presently employed in a job that required extensive interaction with a computer via a video terminal. So, while there were some computer experienced subjects, the majority can be classified as computer inexperienced. It is, therefore, quite reasonable to conclude that, as a group, subjects were relatively computer inexperienced.

Apparatus. The sentences were displayed to subjects on a Volker-Craig, 12 in, VC404 video display terminal, which had a 24 row X 80 column display format, a blinking cursor, and a detachable typewriter style keyboard. The computer program controlling the experiment allowed only the use of the alphanumeric and punctuation keys; the use of other keys, e.g., the return key, resulted in a KEY NOT IN USE error message being printed on the screen.

A custom made cursor control box, with seven keys arranged in two rows, provided the means to control cursor movement on the display screen. The top row consisted of five keys, from left to right: (1) cursor left, move the cursor left one space (2) cursor right, move the cursor right one space (3) insert, insert a space "under" the cursor (4) remove, delete a single character, and (5) next sentence, initiate the next trial. Directly beneath the cursor left key was (6) tab left, move the cursor left ten spaces, and directly beneath the cursor right key was (7) tab right, move the cursor right ten spaces. Each key's function was labeled above it.

The program provided five different user error messages including the previously mentioned KEY NOT IN USE. NOTHING TO DELETE was printed if a subject pressed the remove key when the cursor was to the right of the rightmost, non-blank charac-

ter in a sentence. TYPE FROM KEYBOARD TO LENGTHEN LINE was printed if a subject pressed the insert key when the cursor was to the right of the rightmost, non-blank character in a sentence. END OF LINE FOUND was printed if an attempt was made to move the cursor more than one character past the rightmost character on the screen. FRONT OF LINE FOUND resulted if an attempt was made to move the cursor past the leftmost character of a sentence. Error messages were printed, starting in column one, two lines (rows) beneath the lower sentence on the screen. After the message was written the cursor returned to the position it held before the error condition. The error message was removed from the screen with the next keypress.

Experimental Task. On each trial, a single sentence, printed twice, was displayed on the video screen in front of the subject. The first copy of the sentence was written without errors, while the second copy, written underneath the first, contained errors. The subject's task was to correct the second copy so that it matched the first. Errors were clearly marked by overlining the location of errors in the first copy of each sentence. The same 30 experimental sentences, in the same order, were corrected by each subject following practice with seven practice sentences. When the subject completed the last experimental sentence, he/she opened the

lab door, the experimenter then entered the lab and gave him/her a satisfaction with the system rating scale, and a general information questionnaire to complete (Appendix A).

Practice Sentences. There were seven practice sentences (Appendix B) to be corrected by subjects to familiarize them with the equipment and experimental task. These sentences were composed by the experimenter to be simple, concrete, and unambiguous. Errors were placed in the sentences so that each cursor control key would be practiced more than once. In addition, the general format of the errors was introduced, i.e., all types of errors were represented across the practice sentences, and there were variable numbers of errors in the sentences.

Experimental Sentences. There were 30 experimental sentences, selected from a local newspaper on the basis of the experimenter's judgment that the sentences were simple, concrete, unambiguous out of their original context, and non-controversial. An examination of the experimental sentences (Appendix C) shows that this was largely achieved, and that it is reasonable to assume that the readability of the sentences presented no difficulty for subjects.

Each experimental sentence had at least one error, with the number of errors, the types of error, and the posi-

tioning of errors in a sentence determined randomly in the following manner. First, the 30 sentences were randomly grouped into ten blocks of three sentences per block. Within each block, one, two or three errors were randomly assigned to each sentence for a total of six errors per block, and 60 errors over the 30 sentences.

Second, the type of error was randomly assigned to each error. The types were (1) transposition, (2) extra character, and (3) missing character. Third, one, two or three characters were then randomly assigned to each error by grouping all errors into twenty blocks of three errors per block, and randomly assigning the number of characters in an error to errors in each block. Thus, there were a total of 12 characters involved in the six errors in each block. Over the 30 sentences there were 120 characters involved within the 60 errors.

Finally, sentences were partitioned into six ten character field positions, with errors in a sentence being randomly assigned to a field with the restriction that each field position contain at most one error. The actual position of the error within a field and the characters involved, were arbitrarily determined by the experimenter.

Instructions. Each subject was instructed by the experimenter on the word processing equipment and task, and

each completed seven practice sentences. The purpose of the practice session was to ensure that each subject understood the operation of the system, and practiced using all the function keys. The experimenter used the first practice sentence to demonstrate each cursor control function, and for the second and third practice sentences the experimenter told the subject how to correct the errors. For the remainder of the practice sentences the experimenter assisted only when a subject was apparently having difficulty. After completion of the last practice sentence, the experimenter told each subject, "when I leave, press the next sentence key, and begin working at your own pace." (Instructions Appendix D).

Design. A 3 X 3 factorial design was used with probability of service at three levels, .25, .50 and 1, and minimum delay at three levels, .2, .4 and .8 seconds. These two parameters (probability of service and minimum delay) determined the nature of the delays for the experimental conditions.

Dependent Measures. Subjects rated (1) their overall satisfaction with the word processing system, (2) their satisfaction with the system's speed of responding to commands, and (3) their satisfaction with the ease of using the cursor control commands. A five point Likert-type scale was used for each of the three items (see Appendix E).

Four different performance measures were considered. First, corrected time is the difference between the total time taken to correct the experimental sentences and the total delay time (when the control box functions were inaccessible to the user). Second, the number of errors is the number of mismatches between the first and second copies of experimental sentences, i.e., the number of errors which the subject attempted to correct but which were still incorrect. Third, the number of cursor control keypresses is the number of cursor control box keypresses made correcting the sentences while there was no delay in effect. This measure provides an indication of the relative efficiency with which the cursor control functions were used in correcting the sentences, with fewer keypresses indicating greater efficiency. Fourth, the number of keyboard keypresses made while correcting the sentences was recorded. As was the case for cursor control keypresses, fewer keypresses are taken as indicating greater efficiency.

Other Measured Variables. The total time required to correct the sentences was recorded, as was delay time during which the cursor controls were not operational. Delay time was not considered a dependent variable because it primarily reflects system behavior and not the behavior of subjects. Total time was not considered a dependent variable because it consisted of two components: delay time and corrected time.

Although the number of cursor control keypresses made by subjects during delays was also recorded it was not considered a dependent variable because of the correlation between the opportunity to keypresses and the delay conditions. That is, subjects in conditions with shorter mean delays had less opportunity to keypress during delays. Thus, if subjects in the longer delay conditions keypressed more than those in the shorter delay conditions it would be impossible to determine whether this was a true effect of the manipulations or simply an artifact of the experimental paradigm.

RESULTS

The system response time characteristics obtained for the experimentally generated delays (Table 2) conformed well to those predicted by the model.

Thus, we are assured that the implementation of the model is highly effective, producing predictable results.

In addition, the relationship between the experimental manipulations and the mean and standard deviation of delay should be noted. As the minimum delay in service increases from .2 to .8, the mean delay and standard deviation increase. On the other hand, as the probability of service increases (or the probability of postponement decreases) the mean delay and standard deviation of delay decrease.

Outliers, or extremely deviant performers, can seriously distort parameter estimation when fitting a model to data. A simple technique for handling such outliers is to identify them by examining the data, eliminate their influence in fitting a model to the data, and to separately study their response tendencies and characteristics. Consequently, in this study four outliers were detected and dropped from the analyses. These individuals failed to perform the task as suggested in training and are discussed in Appendix F.

Table 2

Obtained Means and Standard Deviations
On System Response Time

Probability of Service	Minimum Delay			
	.2	.4	.8	
.25	.81 (.8)	1.57 (1.6)	3.12 (3.2)	1.87
.5	.40 (.4)	.80 (.8)	1.58 (1.60)	.93
1.0	.20 (.2)	.40 (.4)	.80 (.8)	.47
	.47	.93	1.87	

Standard Deviations of Delay

Probability of Service	Minimum Delay			
	.2	.4	.8	
.25	.71 (.69)	1.35 (1.39)	2.75 (2.78)	1.62
.5	.28 (.28)	.57 (.57)	1.12 (1.14)	.67
1.0	0 (0)	0 (0)	0 (0)	0
	.32	.65	1.29	

Note: Values in parenthesis are those predicted by the delay generation model.

One measure, the number of errors made correcting the sentences, was also omitted from consideration due to the large number of subjects making no errors. As shown in Table 3, all conditions contained two or more errorless performers (see Appendix G for more details).

Omitting the four outliers, the remainder of the observations were used in a multivariate analysis of variance, and discriminant analyses performed on the 3 X 3 factorial design incorporating the remaining six dependent variables, and utilizing the delay generation parameters - minimum delay and probability of service as independent variables. Since it takes into account the number of dependent measures used in the analysis, the multivariate analysis of variance provides some protection against type I errors. The discriminant analyses provide information about the relationships between the dependent measures. Each discriminant being the weighted sum of the original dependent variables that maximally discriminates the groups.

Similar analyses were done for both satisfaction ratings and performance measures. These analyses were followed by an examination of the relevant structure coefficients. A structure coefficient is the simple correlation between each dependent variable, in a set of dependent variables, with the discriminant function or discriminant that was obtained with each multivariate analysis. In addition, separate univariate

Table 3

Number of Subjects With No Errors
For The Experimental Conditions

Probability of Service	Minimum Delay			Total for Each Probability
	.2	.2	.8	
.25	4	6	5	15
.50	6	2	4	12
1	5	3	3	11
Total for Each Delay	15	11	12	38/

analyses of variance were performed on each dependent measure.

The structure coefficients were used to help determine which variables or sets of variables were highly correlated with group differences and to help define the discriminant substantively (Borgen & Seiling, 1978; Spector, 1980).

Effects of Probability of Service. The three multivariate analyses for probability of service and the univariate F's showed a high degree of correspondence. A single significant discriminant was found for the six measures combined, R^2 ($M = 1.5$, $N = 35$, $S = 2$) = .35, $p < .01$, for the three satisfaction measures R^2 ($M = 0$, $N = 36.5$, $S = 2$) = .13, $p < .05$, and for the three performance measures R^2 ($M = 0$, $N = 36.5$, $S = 2$) = .23, $p < .01$. The fact that only one discriminant was significant indicates a unidimensional basis for the differences.

In general, the group means on the discriminant (the group centroids) change linearly with increases in the probability of service. Centroids were 1.65, .74 and -.06 for probabilities of .25, .5 and 1, respectively, when the six measures were considered jointly. Comparable linear changes were found in the smaller analysis.

Interpretation of the centroids is facilitated by examining the correlation between the discriminant and the individual dependent variables, i.e., the structure coefficients. That is, knowing that there are reliable differences due to the probability of service, it is of interest to find out which dependent variables in the discriminant are reflected in this difference. The structure coefficients for the first discriminant for all three analyses are shown in Table 4. For example, column one presents the structure coefficients for the analysis based on all six dependent variables combined. In this discriminant, overall satisfaction ($r = .49$), speed satisfaction ($r = .81$) and cursor control keypresses ($r = .46$) are most highly correlated with the between group differences suggesting that these measures play an important role in determining the group differences. On the other hand, the small negative correlations for ease satisfaction, keyboard keypresses and corrected time indicate that these variables all decrease slightly or are unrelated to increases in the discriminant. Similar structure coefficients are found in the separate analyses of the satisfaction and performance measures, as shown in the second and third columns of Table 4.

To supplement the information provided by the multivariate analyses, the dependent variables were analysed individually. Univariate analyses of variance summarized in

Table 4

Structure Coefficients For First Discriminant
And Univariate Fs For the Probability
Of Service

Variable	Multivariate Analysis			Univariate F's
	Combined	Satisfaction Ratings	Performance Measures	F
Overall Satisfaction	.49	.59	-	3.29*
Speed Satisfaction	.81	.96	-	10.11**
Ease Satisfaction	-.21	-.24	-	<1
Cursor Control Keypresses	.46	-	.67	2.20
Keyboard Keypress	-.19	-	.23	<1
Corrected Time	-.10.	-	.24	<1

* $p < .05$

** $p < .01$

the last column of Table 4 show that both overall satisfaction and speed satisfaction are statistically significant, while cursor control keypresses fail to reach significance. For all three measures, the group means increase monotonically as probability of service increases.

In summary, considering all of these analyses, it is clear that decreasing the probability of service, thereby increasing both mean delay and variance of delay, results in decreased overall user satisfaction as well as less satisfaction with speed of service. On the other hand, a corresponding decrease in the number of cursor keypresses, which correlates well with the discriminant should be considered marginally significant when the univariate results are considered.

Effect of Minimum Delay. There were reliable differences in the three levels of minimum delay (.2, .4 and .8 sec) on the first discriminant when all six dependent variables were combined, R^2 ($m = 1.5$, $N = 35$, $S = 2$) = .34, $p < .01$; when the satisfaction measures were analysed separately, R^2 ($M = 0$, $N = 36.5$, $S = 2$) = .29, $p < .01$; and when the performance measures were analysed separately, R^2 ($M = 0$, $N = 36.5$, $S = 2$) = .14, $p < .10$. In all three cases, the second discriminant function was not significant. For the three significant discriminant functions, there were linear

decreases in group centroids as the delay increased (e.g., .81, .22 and -.83 and for .2, .4 and .8 sec delays, respectively, when satisfaction and performance measures were included).

As was the case for the main effect of probability of service, speed satisfaction, overall satisfaction and number of cursor control keypresses were the variables most highly correlated with the discriminant. The structure coefficients for the three discriminants are shown in Table 5. From the signs of the structure coefficients, it is clear that speed satisfaction, overall satisfaction and number of cursor control keypresses all decreased as the minimum delay increased. The structure coefficients for both ease satisfaction and keyboard keypresses are larger than those obtained for the probability of service indicating a greater involvement in the differences between delay groups. As delays increased the number of keyboard keypresses and satisfaction with ease of use increased slightly. Corrected time changed minimally or did not change.

The univariate analyses for the individual measures are presented in the last column of Table 5. Differences in speed satisfaction and cursor control keypresses were both significant. Means for overall satisfaction, like the previous two variables, were in the right direction, but the differences were not reliable.

Table 5

Structure Matrix for First Discriminant
For The Minimum Delay

Variable	Multivariate Analysis			Univariate F's
	Combined	Satisfaction Ratings	Performance Measures	F
Overall Satisfaction	-.43	.45	-	1.70
Speed Satisfaction	.83	.88	-	10.58**
Ease Satisfaction	-.40	-.44	-	2.73
Cursor Control Keypresses	.56	-	.88	4.78**
Keyboard Keypress	-.36	-	-.57	2.61
Corrected Time	.14	-	.29	<1

Interaction Between Probability of Service and Minimum Delay. There was no evidence of an interaction between the two independent variables. All three multivariate analyses resulted in non-significant discriminants, largest $p > .05$. In addition, univariate analyses of variance on the six dependent measures all produced non-significant results, largest $F(4,77) = 1.53$, $p < .15$ for the speed satisfaction measure.

DISCUSSION

It was expected that as system response time became worse, satisfaction with the system would decrease. This was indeed found for both decreasing levels of probability of service and increasing levels of minimum delay. In both cases the satisfaction of subjects', as represented by the respective discriminants, decreased as expected. However, the decreases were found mainly for satisfaction with the speed of system response, and, to a lesser extent, for overall satisfaction with the system. Rated ease of using the cursor controls either did not change, or increased slightly with increases in minimum delay and decreases in probability of service. It is not surprising that satisfaction with ease of use did not vary substantially because the complexity of the task was held constant from one condition to the next. It was conceivable that dissatisfaction with one system attribute (e.g. system response time) might be compensated for by satisfaction with another system attribute (e.g. ease of using the cursor controls), or that a single system attribute might be so central to a task or system that all satisfaction measures would be strongly affected by that attribute. This was clearly not the case. It is interesting that subjects produced ratings consistent with the manipulations and that satisfaction is not a unitary concept.

As well as affecting user satisfaction, unfavourable system response time characteristics were also expected to influence subjects' performance (Miller, 1968). That this occurred is evidenced by the linear relationship between the probability of service and its discriminant, and minimum delay and its discriminant. However, like the discriminants for the satisfaction ratings the performance discriminants cannot be defined simply as they reflect some degree of complexity. The increasing values on the discriminants represent increasing use of the keyboard, increasing time taken by subjects to correct the sentences and decreasing use of the cursor controls. Somewhat similar to the situation for the satisfaction discriminants, a single dependent variable, in this case cursor control keypresses, shows the strongest relationship with minimum delay and probability of service. The results indicate that the number of cursor control keypresses required to correct the sentences decreased as the system response time became more unfavourable. Because the cursor control keypresses initiated the delays, it is evident that subjects responded to this unsatisfactory state of affairs by an adjustment in response style. In summarizing possible user responses to an inadequate system, Ramsey and Atwood, 1980 provide a list of eight possible responses. One possibility, partial use of the system, a tendency to avoid unsatisfactory parts of the system even if such adjustments are inefficient, might

account for the small number of cursor control keypresses at greater delays. This tendency is frequently observed in users of low discretion with respect to whether or not they use the system, such as clerks (or students participating in an experiment). There are at least two possible strategies on the part of users that could account for the decreased use of the cursor functions for increased delays. First, subjects could have adopted a strategy favouring the keyboard over the cursor controls, i.e., positioning the cursor within a few characters of the error, and typing over characters to reach the error rather than positioning the cursor exactly. This would account for the increase in the number of keyboard keypresses and possibly the increase in the time required to correct the sentences depending on typing ability. Presumably subjects with poorer typing ability would be at a disadvantage with the self-imposed increase in typing demand. A second possible hypothesis would suggest that subjects in the lower probability of service conditions may have used the cursor controls more judiciously, i.e., taking more care to estimate distances to errors and using the combination of cursor left/right (one space) and tab left/right (ten spaces) keys to minimize "over-shooting" errors with the cursor. Such a strategy would reduce the number of cursor control keypresses needed to correct an error. The extra care taken in using the cursor

functions also suggests an increase in the time required to correct the sentences through increased planning and thinking time. However, this hypothesis does not easily account for the small increase in keyboard keypresses.

The notion that system response time appreciably affects the speed at which users work on a system (Boies, 1974) suggests that subjects in unfavourable delay conditions would tend to have large values on corrected time. This is clearly not supported by the univariate analyses. The system response time conditions had no effect on the time taken to correct the sentences. It is clear that the Boies effect does not generalize to the current task which differs from Boies's situation in a number of ways. It would be of interest to investigate some of the task differences to determine which are critical to the effect. Some differences are as follows. First, the present experiment makes little in the way of information processing demands on the subject since errors are clearly marked, corrected with a small number of subtasks, and requires little short term memory. Miller (1968) certainly would not predict a great deterioration in performance in general under such conditions, although he does not address conditions which would be necessary to produce the pacing effect described by Boies. Second, because the present experiment allows type-ahead, which the Boies study does not

do, the subjects were less paced by the response time. It is possible that the Boies effect does not occur unless users are forced to wait. Finally, Boies based his observations on research in an office environment. This experiment might have failed to replicate the effect because it is only a short term, low pressure and novel situation. These differences suggest a number of directions for future research. First, the information processing load for the subject could be increased by making the task more difficult, such as by requiring the subjects to detect the errors instead of having them marked. Second, the type-ahead feature could be removed and keypress responses could be delayed to determine if the unavoidable nature of the delay is a contributing factor to the Boies effect. Finally, an effort could be made to replicate the experiment under conditions which more closely resemble office working conditions. Time pressure could be incorporated, or users could be asked to work with the system over a longer period of time.

REFERENCE NOTES

1. Tombaugh, J. & Dillon, R.. Videotext: Who uses it and how well? Paper submitted to Eastern Psychological Association.

REFERENCES

- Abrahamson, D.S. & Murray, R.P. Human factors study. Experiments 1 & 2. Keyboard type & system response delay. Report to the Manitoba Telephone System, 1980.
- Boies, S.J. User behavior on an interactive computer system. I.B.M. Systems Journal, 1974, 13, 1-18.
- Borgen, F.H. & Seling, M.J. Uses of discriminant analysis following MANOVA: Multivariate statistics for multivariate purposes. Journal of Applied Psychology, 1978, 63, 689-697.
- Carbonell, J.R., Elkind, J.I. & Nickerson, R.S. On the psychological importance of time in a time-sharing system. Human Factors, 1968, 10, 135-142.
- Doherty, W.J. Human and computer performance. In B. Shackel (Ed.), Man computer communication, (Vol. 1). Maidenhead England: Infotech International, 1979.
- Finn, J.D. & Mattsson, I. Multivariate analysis in educational research: Applications of the multivariate program. Chicago: National Educational Resources, 1978.
- Kerlinger, F.N. & Pedhazur, E.J. Multiple regression in behavioral research. New York: Holt, Rinehard and Winston, 1973.
- Klatzky, R.L. Human memory: Structures and processes. San Francisco.: W.H. Freeman, 1975.
- Miller, R.B. Response time in man-computer conversational transactions, AFIPS Conference Proceedings (Fall Joint Computer Conference 1968). Washington: Thompson Book Company, 1968.
- Ramsey, H.R. & Atwood, M.E. Human factors in computer systems: A review of the literature. JSAS Catalogue of Selected Documents in Psychology, 1979, 10(2), 49. (Ms. No. 2039)

Spector, P.E. Redundancy and dimensionality as determinants of data analytic strategies in multivariate analysis of variance. Journal of Applied Psychology, 1980, 65, 237-239.

APPENDIX A

General Information Questionnaire

Below you are asked to provide some personal information that might help us interpret the results of this experiment. The information you provide will be kept strictly confidential.

1. Sex:
2. Age:
3. Major Subject:
4. Number of Computing Science Courses:
5. Approximately how many hours have you spent using a computer via a teletype or video screen? _____

Comments

The space provided below is for any observations or comments you may have concerning this experiment.

APPENDIX B
Practice Sentences

v

Sentence 1

The cāt is in the tree.

The cet is in the tree.

2

This is the wāy out.

This is the wy out.

3

Thēy paid toō much mōney.

Thay paid tooo much money.

4

It was a long and hard dāy at work.

It wass a long ad hard daay at work.

5

The smāll car hit the huge truck.

The smull car hit the huge truck.

6

Thē cat atē the small bīrd.

Te cat atee the small bord.

7

The dog chāsed the cat up the tree.

The dog chsed the cat upp the tree.

APPENDIX C

Experimental SentencesSentence 1

One shop in the town was bombed but no one was hurt.

* One shope in the town was bombed but no one was hurt.

2

Instead, it would find and develop its own oil and gas.

Instead, it would find and velop its own oilll and gas.

3

The answer to the first two questions is a firm "yes".

The aner to the first tow quest is a firm "yes".

4

The rooms are, almost invariably, all occupied.

The rooms are, alsmot invariably, all occupied.

5

There's a resistance in lawyers to deal with joint custody.

There's resistance inn lawyers to deal with jtnoi custody.

6

Before being announced to skate, he fell and later
skated off the ice.

Before being announced to se, he fell and lat
skated off the ice.

7

Two hamburgērs, two frēnch fries, a milkshake and a coffee.

Two hamburgres, two frch fries, a milkshake and a coffee.

8

Now, all of these ideas and proposals appear to have mērit.

Now, all of these ideas and proposals appear to have mt.

9

Is the wintēr a good tīme for sūch a trip?

Is the wintre a good tie for shuc a trip?

10

Fighting, she said, would "lead straight to nūclēar war."

Fighting, she said, would "lead straight to nuelēear war."

11

"It was fun ānd I'd do it again evēn if they didn't pāy
me," she said.

"It was fun I'd do it again evn if they didn't paiye
me," she said.

12

When children grōw up and move away there is nobody to
replāce them.

When childrēn gwro up and move away there is nobody to
replae them.

13

The cause of both fires is still under investigation.

The causes of both fi is still udenr investigation.

14

This year it looks like the big fish were "the ones that got away."

This year it looks like the big fish were "the osne that got away."

15

That was in 1975, and they haven't looked back since.

That was in 1759, and they haven't loked back since.

16

Bed and breakfast is a unique British institution.

Bed andand breakfast is a unique Brith institution.

17

Only recently has any actual research and data-gathering begun.

Only recently has any actal researche and data-gathering ubegn.

18

He liked what he heard and invited the group to Nashville.

He lkeid what he heard and invited the group to Nashville.

19

He said the contracts were being picked up by the players
new clubs.

He said the contracts were being picked up by the players
new clubs.

20

Skater's injuries ranged from minor bruises to broken legs.

Skater's injuries ranged from minor bruises to broken legs.

21

Some firms are having more problems than others.

Some firms are having more problems than others.

22

Children's school pictures lined the bookcase.

Children's school pictures lined the bookcase.

23

All five men arrived at the jail.

All five men arrived at the jail.

24

When a stroke occurs the blood flow is blocked, killing
the cells.

When a stroke occurs the blood flow is blocked,
killing the cells.

25

A Swiss member of her expedition did die of exhaustion and exposure.

A Swiss member of her expedition did die of exhaustion and exposure.

26

Cooking oil solidifies after it cools and tends to collect in sewers.

Cooking oil solidifies after it cools and tends to collect in sewers.

27

The Equal Rights Amendment is in deep trouble.

The Equal Rights Amendment is in deep trouble.

28

The workshop will emphasize the importance of being positive.

The workshop will emphasize the importance of being positive.

29

Indeed, they are prohibited by law from voting.

Indeed, they are prohibited by law from voting.

30

Accommodation in London itself is difficult and expensive.

Accommodation in London itself is difficult and expensive.

APPENDIX D

Experimental Instructions

The experimenter met the subject outside the lab, introduced himself, and showed the subject into the lab seating him or her facing the CRT. The experimenter allowed the subject time to settle and insure that any books, etc. belonging to the subject were placed out of the way. The experimenter then began the experimental instructions with a general explanation of the purpose of the experiment, followed by a demonstration of the cursor control box and finally the seven practice sentences. The experimenter used the following script as a guide in delivering the instructions.

Before we start you on the proofreading task, I first want to give you a general idea of what the study is about. We are interested in finding out what the best characteristics are for a computerized word processing system. To do this we are having people do a proofreading task on computer systems that have different characteristics, and by having them rate the systems they use, we will get an indication of what the best characteristics are for a word processing system. What you will be doing is proofreading and

correcting sentences using a computerized word processing system. The sentences will be presented to you on the video screen, and you will correct each one as indicated using the editing system of the word processor.

Any questions?

What I'll do now is give you a short demonstration of what each key does (points to cursor control box) and then we will do seven practice sentences together.

At this time the demonstration sentence was on the CRT screen:

I plan to take physics next semester.

I plan to take physics next semester.

The sentences you will be correcting will be presented in the same format as the sentence on the screen right now. This box on your left (points) is the cursor control box which you will use to position the cursor to correct the sentences. The cursor control box has seven different commands (as demonstrating each key):

- (1) a cursor left key to move the cursor one space to the left.

- (2) a cursor right key to move the cursor one space to the right.
- (3) a tab left key to move the cursor ten spaces to the left.
- (4) a tab right key to move the cursor ten spaces to the right.
- (5) an insert key to insert a space between two characters.
- (6) a remove key to remove a character and which adjusts the remaining characters by shifting them one space to the left.
- (7) and a next sentence key which will present the next sentence after you have finished correcting the current sentence.

Any questions?

What we'll do now is go through seven practice sentences together so that you'll know how to use the editing system. It is important that you are clear on how to use the system properly since I won't be here while you are doing the experiment. If you have any questions make sure you ask them before I leave.

The experimenter then pressed the next sentence key to bring up the first of seven practice sentences.

Practice Sentence One

The cāt is in the tree.

The cet is in the tree.

Notice that the sentence on top has no errors in it, while underneath is the same sentence with an error (pointing). The line above the first sentence marks the approximate location of the error in the second. The first thing you have to do to correct the sentence is to move the cursor to the error. You can do this by using the tab left key and the cursor left key. O.K., do that now. To correct the error just type over the incorrect letter. Now that the error is corrected you want the next sentence. To get it, press the next sentence key once.

Two

This is the wāy out.

This is the wy out.

Position the cursor so that you'll be able to correct the error, and correct it using the insert key. The insert key will create a space between

the "w" and "y" so that you can type the missing letter. Get the next sentence.

Three

Thēy paid too much money.

Thay paid tooo much money.

O.K. This sentence has two errors that you have to correct.

From this point the subject was allowed to continue with the remainder of the practice sentences with the experimenter intervening only if the subject was apparently having difficulty. In such cases the experimenter prompted the subject on how to correct the error. Also, if the subject pressed the next sentence key before all the errors were corrected, this was mentioned by the experimenter. If during the practice session the subject inquired about a system delay the experimenter always responded "sometimes it takes a while for the computer to do what you want it to".

APPENDIX E

Satisfaction Rating ScaleWord Processing System Rating Scale

Below are some statements concerning the word processing system. Please indicate your response by circling the appropriate number.

RATE:

(a) Your overall satisfaction with the performance of the word processing system.

5	4	3	2	1
extremely	very	satisfied	somewhat	not satisfied
satisfied	satisfied		satisfied	at all

(b) The speed with which the word processing system responded to your commands.

5	4	3	2	1
extremely	very	satisfactory	somewhat	not satisfac-
satisfactory	satisfactory		satisfactory	tory at all

(c) The ease of using the word processing commands.

5	4	3	2	1
extremely	very	easy	somewhat	not easy
easy	easy		easy	at all

APPENDIX F

Outliers

There was a problem with outliers in the data. A cursory examination of the raw data (Appendix L) shows four different outliers on two different dependent measures - cursor control keypresses and keyboard keypresses - as well as on the additional measured variable, the number of cursor control keypresses made during delays. To more formally assess the extent to which each outlier deviated from its own experimental group its standard z score was calculated by first dropping each outlier from its group, and then using the remainder of the subjects in that group to calculate the group means and standard deviations used to calculate the z scores. Subject (number) 51 was an outlier on two variables, cursor control keypresses and keyboard keypresses with respective standard scores of $z = 6.06$ and $z = 16.04$. Subject 71 had a standard score of $z = 9.16$ on keyboard keypresses, while subject 16 had a standard score of $z = 29.58$ on delay cursor control keypresses. Finally, subject 58's standard score on keyboard keypresses was $z = 46.84$. Clearly these subjects were not treating the experimental situation in the same way as other subjects, and clearly they must have done the task in some rather unconventional ways, e.g., for some errors subject retyped large sections of the sentence rather than

moving the cursor to the error for correction. Obviously, the outlying subjects did the task in a manner inappropriate for this experiment and it would therefore be misleading to include them in the analyses.

The elimination of outliers from the data (Table 6) results in non-orthogonality (Table 7) among the various effects. The problem correlations are those between vectors coding different effects since these correlations can make ambiguous the partitioning of the variance for the different independent variables. However, only a small number of subjects from different conditions were involved, and consequently these correlations were small, none exceeding .04. Because of the non-orthogonality the classical experimental approach was used for both univariate (Kerlinger, 1973) and multivariate (Finn & Mattsson, 1978) analyses.

Table 6
Cell Frequencies With Outliers Removed

		<u>Minimum Delay</u>		
		.2 sec.	.4	.8
Probability of Service	.25	9	10	9
	.50	10	10	9
	1	10	9	10

Table 7

Intercorrelations of Vectors Coded
for Factorial Design: Outliers Excluded

Vector

	1	2	3	4	5	6	7
1							
2	.4911						
3	-.0002	.0173					
4	-.0353	-.0002	.4911				
5	.0000	.0000	.0430	.0215			
6	.0005	.0220	.0220	.0005	.5003		
7	-.0003	.0215	.0215	-.0003	.4801	.2408	
8	.0215	.0430	.0000	.0000	.2368	.5003	.4801

Note: Vectors 1 and 2 code probability of service.

Vectors 3 and 4 code minimum delay.

Vectors 5 through 8 code the probability X delay interaction.

APPENDIX G

Number of Errors Made By Subjects

Since subjects were required to match the second copy of each sentence there existed the possibility that they might leave some errors uncorrected. That is, the delay manipulations might have interfered with whatever psychological processes were involved in correcting the sentences. However, an examination of the number of errors made by each subject in each condition showed that a few subjects in each condition contributed the majority of the errors. For example, in condition one (see Table 8 for the assignment of number labels to conditions) Subject 22 accounted for 34% of the total number of errors for his/her group. Likewise in condition three, Subject 71 accounted for 57% of that group's errors, Subject 55 accounted for 35% of group five's errors and Subject 33 accounted for 86% of his/her group's errors. Also, in group seven Subject 30 contributed 50% to that group's total number of errors. It does not seem that a statistical analysis concerning group differences would achieve very much, since certain individuals contribute disproportionately to their respective group totals. In fact, with the benefit of hindsight it could be argued that errors was a poor choice for a dependent measure. The task was well defined, and there was no appreciable memory load on a subject. Further since subjects worked at their own

pace, they had ample time (if they so chose) to carefully check each sentence before starting the next. Simply said, the task did not demand enough of subjects. The means and standard deviations on the number of errors are given in Table 8,

Table 8
Means and Standard Deviations
On the Number of Errors

		<u>Minimum Delay</u>		
		.2 sec.	.4	.8
Probability of Service	.25	2.2 (3.15) 1	1.3 (2.0) 2	1.44 (2.19) 3
	.50	1.3 (1.95) 4	1.41 (.94) 5	2.56 (6.21) 6
	1	1.8 (2.86) 7	2.56 (2.46) 8	2.10 (2.02) 9

Note: Standard deviations are in brackets.

Condition number is in lower left corner of each cell.

APPENDIX H

Univariate Analyses

The final step in the data analysis was univariate analyses of variance on each dependent measure. The extent to which the univariate analyses agree with the multivariate, e.g., speed satisfaction showing large effects and cursor control keypresses being prominent among the performance measures, we can be confident that we have isolated the sources of the differences between groups. In addition, these analyses will allow closer examination of the relationships between each dependent variable and the independent variables. The results for the satisfaction ratings are given first followed by those for the performance measures.

Since system response time characteristics should figure into a users overall evaluation of the system, it would be reasonable to expect that manipulations of system response time characteristics would affect users ratings of overall satisfaction. The analysis of overall satisfaction did show a significant effect for the probability of service, $F(2,77) = 3.29, p < .05$, with ratings of overall satisfaction increasing with increases in the probability a delay would end (Table 9). Only the linear component for this trend was significant, $F(1,77) = 5.65, p < .05$, while for the quadratic component, $F < 1$. The effect for minimum

Table 9

Means on Overall Satisfaction for
Probability of Service

Probability	Mean	Standard Error
.25	2.96	.19
.50	3.41	.20
1	3.62	.16

delay, and the interaction, probability X delay, were not significant, in order, $F(2,77) = 1.74$, $p > .05$ and $F < 1$. Given the multivariate results showing the relatively small structure coefficients for overall satisfaction (Table 4) and the relatively small univariate effect (Table 32), the significance of the effect should be treated with some caution. However, the results do seem to tentatively support the conclusion that subjects were sensitive to the manipulation of system response time characteristics.

The experimental manipulations directly affected the speed with which the system responded to a user's cursor control commands. Given that these manipulations seemed to influence ratings of overall satisfaction, and further that speed satisfaction had very large structure coefficients for all the discriminants which included satisfaction ratings, a large effect on speed satisfaction ratings would be expected. This was in fact the case. The speed satisfaction ratings showed a significant effect for both the probability of service and minimum delay, in order, $F(2,77) = 10.11$, $p < .01$, and $F(2,77) = 10.58$, $p < .01$. The mean ratings increased with the probability of service, and decreases in the minimum delay (Table 10,11). In both cases only the linear components of the trends were significant, for the probability of service and minimum delay, in order,

Table 10

Means on Speed Satisfaction for
Probability of Service

Probability	Mean	Standard Error
.25	2.03	.17
.50	2.65	.24
1	3.24	.22

Table 11

Means on Speed Satisfaction for
Minimum Delay

Delay	Mean	Standard Error
.2 sec.	3.17	.14
.4	2.79	.24
.8	1.96	.24

$F(1,77) = 19.55, p < .01$ and $F(1,77) = 21.07, p < .01$ while for the nonsignificant quadratic components, both F 's < 1 . The interaction, probability X delay was not significant, $F(4,77) = 1.529, p > .05$. From these results it is possible to conclude that subjects were sensitive to the experimental manipulations as reflected in the lower mean satisfaction ratings as delay conditions became more unfavourable, i.e., as the probability of service decreased or the minimum delay increased. Indeed the effects appear quite strong (Tables 33, 34).

Finally, since the response time manipulations could not directly affect the ease of using the cursor controls one might expect that ease satisfaction ratings would be similarly unaffected. The multivariate analyses seem to reflect this proposition, with ease satisfaction having generally modest structure coefficients. The univariate results are also consistent with the expectation of no effect. Neither main effect, the probability of service, minimum delay, nor the interaction probability X delay were significant; all F 's < 1 .

Performance Measures. The corrected time was the average of the differences between the total time (Table 12) it took a subject to complete the experimental task, and the delay time (Table 13), during which the cursor control functions

Table 13

Means on Delay TimeProbability of
Service

Minimum Delay

	.2	.4	.8	
.25	4159.33 (216.24)	7694.80 (451.89)	14391.11 (1012.25)	8710.79 (877.92)
.50	2247.20 (141.08)	4131.20 (241.98)	7596.44 (454.53)	4556.96 (444.24)
1.0	1159.60 (46.86)	2078.67 (102.00)	4027.20 (215.86)	2433.65 (242.88)
	2465.59 (244.91)	4723.03 (469.71)	8505.71 (900.17)	

Note: Number in parenthesis is
standard error.

Table 12

Means on Total Time

Probability of Service	Minimum Delay			Total Time
	.2	.4	.8	
.25	18372 (1301.76)	22116.70 (1544.36)	28002.11 (2511.83)	21627.43 (1302.36)
.5	16723.20 (1162.76)	18452.30 (1330.99)	20570.56 (1251.81)	18581.69 (1046.3)
1.0	15085 (1062.58)	14797.67 (1903.34)	16841.40 (1107.31)	16670 (698.81)
	15601.48 (782.08)	18513.45 (756.03)	22804.78 (1268.05)	

Note: Number in parenthesis is
standard error.

Table 14

Means on Corrected TimeProbability of
Service

Minimum Delay

	.2	.4	.8	
.25	14212.67 (1183.54)	14421.90 (1237.08)	13611.00 (1531.81)	14094 (737.18)
.50	14476 (1036.71)	14321.1 (1216.39)	12974.11 (875.29)	13956.48 (605.57)
1.0	13925.4 (1048.09)	12719 (1839.26)	12814.2 (931.70)	13167.83 (727.17)
	14204.41 (605.77)	13858.65 (809.02)	13121.71 (635.72)	

Note: Number in parenthesis is
standard error.

were not accessible to a subject. Thus total time can be thought to be composed of two components: a component attributable to system behavior, the delay time, and a component attributable to user behavior, the corrected time. For total time both main effects were significant, $F(2,77) = 17.45$, $p < .01$, and $F(2,77) = 8.36$, $p < .01$ for the probability of service and minimum delay respectively; while their interaction was not significant $F(4,77) = 1.84$, $p > .05$.

For delay time, the system component of total time, the main effects for both probability of service and minimum delay were significant, in order, $F(2,77) = 182.88$, $p < .01$ and $F(2,77) = 167.30$, $p < .01$. In addition the interaction, probability X delay was also significant $F(4,77) = 21.28$, $p < .01$. Thus it can be concluded that there are substantial differences in system behavior between conditions as reflected by delay time. As stated, the variable total time consists of a component due to the experimental set-up (delay time) and a component due to subjects' behavior (corrected time). There were no significant differences on corrected time for either the probability of service, or minimum delay, both $F's < 1$. The interaction, probability X delay was also not significant, $F < 1$. Thus, it is apparent that the differences between the experimental groups on total time are primarily due to system behavior (delay time) rather than user behavior (corrected time).

Although the experimental task did not require subjects to do much typing, delay conditions could have had an effect on the number of keyboard keypresses required to correct the experimental sentences. For instance since the cursor control box was the locus of system delays, subjects could have avoided use of the control box by typing over characters to an error, rather than positioning the cursor exactly to correct an error, and thereby increasing the number of keyboard keypresses. However, despite this possibility, there was no effect for the probability of service, $F < 1$, although there was a marginally significant effect for minimum delay $F(2,77) = 2.61, p < .08$. The number of keyboard keypresses increasing with increases in the minimum delay (Table 15). The interaction, probability X delay, was not significant, $F < 1$.

Since delays were associated with the cursor controls, the number of keypresses would seem especially likely to be affected by delay conditions. There was in fact a significant effect for minimum delay, $F(2,77) = 4.78, p < .01$, with (Table 16) the mean number of cursor control keypresses decreasing with increases in the minimum delay. The linear component for this trend was significant, $F(1,77) = 8.56, p < .01$, while the quadratic component was not $F(1,77) = 1.00, p > .05$. However, the effect for probability of service was not significant, $F(2,77) = 2.20, p < .12$, nor was the interaction, probability X delay $F < 1$.

Table 15

Means on Keyboard Keypresses for
Minimum Delay

Delay	Mean	Standard Error
.2 sec.	130.55	8.93
.4	125.62	5.46
.8	161.68	18.58

Table 16

Means on Cursor Control Keypresses for
Minimum Delay

Delay	Mean	Standard Error
.2 sec.	551.93	17.11
.4	506.45	15.77
.8	481.11	15.61

APPENDIX I

Tables of Discriminant Weights

The raw score and standard score weights are provided for the first discriminant on each analysis.

Table 17

Analysis with Combined Performance
Measures and Satisfaction Ratings
Probability of Service

Variable	Raw Score Weight	Standardized Score Weight
Overall Satisfaction	.05	.05
Speed Satisfaction	.89	.91
Ease Satisfaction	-.58	-.50
Cursor Control Keypresses	.01	.62
Keyboard Keypresses	.00	.27
Corrected Time	.00	-.96

Table 18

Analysis with Combined Performance
Measures and Satisfaction Ratings
Minimum Delay

Variable	Raw Score Weight	Standardized Score Weight
Overall Satisfaction	-.06	-.06
Speed Satisfaction	.88	.90
Ease Satisfaction	-.68	-.59
Cursor Control Keypresses	.01	.47
Keyboard Keypresses	.00	-.06
Corrected Time	.00	-.50

Table 19

Analysis with Combined Performance
Measures and Satisfaction Ratings
Probability X Delay

Variable	Raw Score Weight	Standardized Score Weight
Overall Satisfaction	.26	.26
Speed Satisfaction	-1.03	-1.05
Ease Satisfaction	-.12	-.10
Cursor Control Keypresses	.00	-.18
Keyboard Keypresses	-.01	-.64
Corrected Time	.00	.46

Table 20

Analysis with Satisfaction Ratings
Probability of Service and Minimum Delay

Probability of Service

Variable	Raw Score Weight	Standardized Score Weight
• Overall Satisfaction	-.04	-.04
Speed Satisfaction	-.96	-.98
Ease Satisfaction	.42	.36

Minimum Delay

Overall Satisfaction	.06	.06
Speed Satisfaction	-.94	-.96
Ease Satisfaction	.67	.58

Table 21
Analysis with Satisfaction Ratings
Probability X Delay

Variable	Raw Score Weight	Standardized Score Weight
Overall Satisfaction	.31	.31
Speed Satisfaction	-1.03	-1.06
Ease Satisfaction	-.39	-.34

Table 22

Analysis with Performance Measures
Probability of Service and Minimum Delay

Probability of Service

Variable	Raw Score Weight	Standardized Score Weight
Cursor Control Keypresses	-.01	-1.22
Keyboard Keypresses	.00	-.09
Corrected Time	.00	1.01

Minimum Delay

Cursor Control Keypresses	.01	.91
Keyboard Keypresses	-.01	-.48
Corrected Time	.00	-.12

Table 23
Analysis with Performance Measures
Probability X Delay

Variable	Raw Score Weight	Standardized Score Weight
Cursor Control Keypresses	.00	-.20
Keyboard Keypresses	-.02	-1.01
Corrected Time	.00	.66

APPENDIX J

ANOVA Summary Tables

Table 24
Overall Satisfaction

Source	DF	MS	F	P
Probability of Service	2	3.21	3.29	.04
Minimum Delay	2	1.70	1.74	.18
Probability X Delay	4	.57	.59	.67
Error	77	.98		
Total	85	1.03		

Table 25

Speed Satisfaction

Source	DF	MS	F	p
Probability of Service	2	10.56	10.11	.00
Minimum Delay	2	11.04	10.58	.00
Probability X Delay	4	1.60	1.53	.20
Error	77	1.04		
Total	85	1.52		

Table 26

Ease Satisfaction

Source	DF	MS	F	p
Probability of Service	2	.48	.64	.53
Minimum Delay	2	2.04	2.73	.07
Probability X Delay	4	.26	.35	.84
Error	77	.75		
Total	85	.74		

Table 27

Total Time

Source	DF	MS	F	p
Probability of Service	2	.37 E09	17.44	.00
Minimum Delay	2	.18 E09	8.36	.00
Probability X Delay	4	39698112	1.84	.13
Error	77	21528352		
Total	85	34389616		

Table 28

Delay Time

Source	DF	MS	F	P
Probability of Service	2	.37 E09	17.45	.00
Minimum Delay	2	.18 E09	8.36	.00
Probability X Delay	4	39698112	1.84	.13
Error	77	21528352		
Total	85	34389616		

2 2

OF / DE

A

Table 29

Corrected Time

Source	DF	MS	F	p
Probability of Service	2	6942428	.48	.62
Minimum Delay	2	8444824	.58	.56
Probability X Delay	4	1982129	.14	.97
Error	77	14456562		
Total	85	13556778		

Table 30

Keyboard Keypresses

Source	DF	MS	F	p
Probability of Service	2	3129.45	.73	.48
Minimum Delay	2	12208.14	2.61	.08
Probability X Delay	4	3402.38	.79	.53
Error	77	4289.26		
Total	85	4374.16		

Table 31

Cursor Control Keypresses

Source	DF	MS	F	p
Probability of Service	2	16775.57	2.20	.12
Minimum Delay	2	36381.06	4.78	.01
Probability X Delay	4	1079.28	.14	.97
Error	77	7611.65		
Total	85	8207.28		

APPENDIX K

Trend Analyses Summary Tables

Note: P = Probability of Service
PS = Probability of Service Squared
M = Minimum Delay
MS = Minimum Delay Squared

Error term for each analysis was from the corresponding ANOVA. Incremental proportions of variance added by a variable were tested. For the non-orthogonal design the main effect not being tested was always entered first.

Table 32

Trend Analysis on Overall Satisfaction for
Probability of Service

Variable	Multiple R	R Square Change	D.F.	F	p
M	.20	.04	-	-	-
MS	.20	.00	-	-	-
P	.32	.06	1	5.65	.02
PS	.34	.01	1	.93	.34

Table 33

Trend Analysis on Speed Satisfaction for
Probability of Service

Variable	Multiple R	R Square Change	D.F.	F	p
M	.41	.17	-	-	-
MS	.41	.00	-	-	-
P	.57	.16	1	19.55	.00
PS	.57	.01	1	.68	.41

Table 34
Trend Analysis on Speed Satisfaction for
Minimum Delay

Variable	Multiple R	R Square Change	D.F.	F	P
P	.39	.15	-	-	-
PS	.40	.01	-	-	-
M	.57	.17	1	21.07	.00
MS	.57	.00	1	.07	.79

Table 35
Trend Analysis on Keyboard Keypresses
for Minimum Delay

Variable	Multiple R	R Square Change	D.F.,	F	p
P	.09	.01	-	-	.5
PS	.12	.01	-	-	-
M	.25	.05	1	5.13	.03
MS	.27	.01	1	1.03	.31

Table 36

Trend Analysis on Cursor Control Keypresses
for Minimum Delay

Variable	Multiple R	R Square Change	D.F.	F	p
P	.21	.05	-	-	-
PS	.22	.01	-	-	-
M	.38	.09	1	8.56	.00
MS	.39	.01	1	1.00	.32

APPENDIX L

Raw DataKey:

SU	Subject number
P	Probability of service
M	Minimum delay
O	Overall satisfaction
S	Speed satisfaction
E	Ease satisfaction
KK	Keyboard Keypresses
CK	Cursor control Keypresses
TT	Total time
DT	Delay time
CT	Corrected time
NE	Number of errors

SU	P	M	O	S	E	KK	CK	TT	DT	CT	VE
13	1	1	4	5	3	84	603	13294	1206	12088	1
21	1	1	3	3	4	216	578	18428	1156	17272	4
30	1	1	3	5	2	85	608	22209	1216	20993	9
40	1	1	5	4	4	162	586	16648	1172	15476	0
48	1	1	4	3	5	107	702	14812	1404	13408	0
66	1	1	4	4	5	91	588	12717	1176	11541	0
79	1	1	4	5	4	74	551	11963	1102	10861	0
85	1	1	3	4	5	185	623	16320	1246	15074	2
03	1	1	4	4	5	81	551	11606	1102	10504	2
59	1	1	4	3	3	123	408	12853	816	12037	0
04	1	2	3	3	4	84	528	24418	2112	22306	0
10	1	2	4	3	4	132	531	14065	2124	11941	2
58	1	2	5	5	4	1481	674	37165	2696	34469	11
36	1	2	4	4	5	168	448	14066	1792	12274	7
37	1	2	4	4	3	173	685	24101	2740	21361	4
49	1	2	5	5	4	143	480	11937	1920	10017	0
67	1	2	2	1	3	131	503	8879	2012	6867	3
81	1	2	2	2	4	146	442	9064	1768	7296	0
23	1	2	5	5	4	100	585	12985	2340	12751	5
90	1	2	5	4	5	141	475	13664	1900	11764	2
06	1	3	4	3	5	180	483	19193	3864	15329	0
15	1	3	4	2	4	78	509	17021	4072	12949	5
24	1	3	2	2	3	113	662	22588	5296	17292	4
28	1	3	3	2	5	90	592	18250	4736	13514	1
39	1	3	3	2	4	135	437	16167	3496	12671	5

SU	P	M	O	S	E	KK	CK	TT	DT	CT	NE
54	1	3	3	3	5	119	395	10956	3160	7796	0
56	1	3	4	1	4	277	542	19267	4336	14931	2
68	1	3	3	2	5	112	409	14849	3272	11577	0
73	1	3	3	3	4	117	450	12020	3600	8420	1
87	1	3	4	3	4	87	555	18103	4440	13663	3
09	2	1	3	2	3	145	459	15350	1920	13430	1
11	2	1	5	3	4	156	531	13546	2020	11526	0
19	2	1	5	5	5	90	665	19922	2654	17268	0
32	2	1	4	3	4	133	438	13726	1734	11992	0
42	2	1	2	1	4	123	476	11159	1854	9302	0
52	2	1	5	4	5	217	543	15475	2166	13309	5
62	2	1	2	3	2	139	787	23926	3170	20756	4
64	2	1	4	4	5	145	489	18003	1994	16009	3
80	2	1	4	2	4	82	573	17406	2384	15022	0
83	2	1	4	5	5	100	645	18719	2576	16143	0
05	2	2	4	3	5	120	550	15387	4476	10911	3
17	2	2	4	1	3	173	432	14061	3528	10533	4
20	2	2	4	4	5	100	591	19693	4792	14901	0
34	2	2	2	3	2	161	495	23835	4004	19831	1
43	2	2	3	3	5	108	421	12003	3316	8687	2
46	2	2	2	2	5	122	374	15296	3084	12212	2
55	2	2	3	2	5	133	718	22421	5720	16701	10
72	2	2	4	3	5	82	543	17065	4292	12773	2
78	2	2	5	5	4	156	507	22652	4120	18532	4
86	2	2	4	4	4	140	489	22110	3980	18130	0

SU	P	M	O	S	E	KK	CK	TT	DT	CT	NE
08	2	3	3	1	5	195	451	18523	7120	11403	0
16	2	3	2	1	3	153	444	17125	7296	9829	1
33	2	3	4	3	3	172	633	25677	9928	15749	19
41	2	3	4	3	5	137	484	20711	7704	13007	1
50	2	3	2	1	4	93	587	27063	9616	17447	0
65	2	3	2	1	5	496	378	20316	6064	14252	0
77	2	3	3	1	5	107	479	20477	7680	12797	0
84	2	3	4	2	5	113	466	19744	7392	12352	1
57	2	3	1	1	5	211	416	15016	6512	8504	0
27	2	3	3	2	4	147	416	17608	6352	11256	2
07	3	1	3	1	2	172	617	25077	4770	20307	5
12	3	1	2	1	4	122	430	15003	3562	11441	0
22	3	1	4	4	5	273	430	18485	3476	15009	9
31	3	1	2	2	3	103	617	21217	5104	16113	1
38	3	1	2	2	4	167	533	16376	4394	11982	0
61	3	1	4	3	4	103	465	22632	4044	18588	4
69	3	1	4	3	4	122	443	15279	3668	11611	0
75	3	1	2	1	5	96	453	13298	3514	9784	0
82	3	1	4	3	4	90	614	17981	4902	13079	1
51	3	1	2	2	2	1075	1022	47704	8048	39656	6
26	3	2	2	2	3	155	679	31639	10636	21003	6
35	3	2	3	2	4	107	439	20781	7200	13581	0
44	3	2	3	1	5	161	444	15267	7032	8235	2
47	3	2	3	3	4	75	540	20056	8304	11752	0
60	3	2	2	1	3	120	452	25574	7200	18374	2

SU	P	M	O	S	E	KK	CK	TT	DT	CT	NE
70	3	2	4	3	5	109	551	25509	9324	16185	0
74	3	2	3	1	5	117	432	20494	6448	14046	0
88	3	2	3	2	4	75	523	21954	8092	13862	3
01	3	2	2	1	3	120	396	15734	5868	9866	0
18	3	2	4	4	5	91	434	24159	6844	17315	0
14	3	3	1	2	5	188	407	24146	12504	11642	0
25	3	3	2	1	4	273	433	27755	13840	13915	4
29	3	3	3	2	5	96	419	20574	11936	8638	0
45	3	3	2	2	4	124	388	22663	11672	10991	0
63	3	3	2	1	5	116	423	24860	13088	11772	0
71	3	3	2	1	4	1160	429	33026	12856	20170	16
76	3	3	4	2	5	416	651	46222	21288	24934	1
89	3	3	4	2	5	123	412	26126	13192	12934	0
53	3	3	4	2	5	126	471	28501	15448	13053	6
02	3	3	5	3	5	86	523	31172	16552	14620	1

END

3 1 H 0 8 H 8 2

FIN