

# Computer Response Time and User Performance During Data Entry

By T. W. BUTLER\*

(Manuscript received October 1, 1982)

In this experiment, subjects entered data at a computer terminal while the response time of the computer was varied systematically. Long average response times were found to be associated with significantly longer subject "think times", as was an increase in the variability of the computer's response time. Six subjects entered five-character letter groups under ten different computer response time conditions. The computer response time distributions used had mean values of 2, 4, 8, 16, and 23 seconds, with two different levels of variability at each mean value. User error rate and user typing time were not significantly affected by computer response time, but computer response time was significantly related to user response time (or think time) with this task. User response time increased slowly and gradually with increases in computer response over the range of stimulus values used. Increases in computer response time variability also increased user response time.

## I. INTRODUCTION

Computer system designers face a serious problem when setting response time requirements for systems that they build. They must confront the competing pressures of providing the most responsive user service possible and of providing this service at the lowest possible cost. Often, cost pressures prevail and the result is a slow, heavily loaded system. Few data are available regarding the effect computer response time has on the users of a system.

Everyone who has used an interactive computer terminal knows that the response time of a system is important for its users. Indeed,

---

\* AT&T Bell Laboratories.

---

Copyright © 1984 AT&T. Photo reproduction for noncommercial use is permitted without payment of royalty provided that each reproduction is done without alteration and that the Journal reference and copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free by computer-based and other information-service systems without further permission. Permission to reproduce or republish any other portion of this paper must be obtained from the Editor.

users may find response time to be a computer system's most salient general characteristic. Nevertheless, there is no firm empirical basis for setting standards of acceptable response times for computer systems. The problem is one both of user performance and of user attitudes and perceptions. The relations of these variables to computer system response time have not been established.

Many papers have noted the importance of Computer System Response Time (CSRT) for users,<sup>1-4</sup> but very few have presented any solid experimental data relating to the problem. Several issues and hypotheses have been put forward in this body of literature:

1. User performance may decline at very long (or very short) computer response times.
2. Variability of computer response time may be as important for users as average response time duration.
3. Different user commands have different response time requirements, depending on the cognitive load that a command imposes on users.

Most early studies of this problem failed to produce meaningful, useful results for one of several reasons. Many studies simply surveyed computer use without systematically manipulating computer response time.<sup>5,6</sup> Subjects in these studies were in a continuously changing environment where most of the variables of interest (e.g., mean response time, response time variability, tasks being performed) were entirely out of the experimenters' control. Other studies have used unrealistic stimulus conditions that confound their results and make their application to this problem unclear. For example, Grignetti and Miller<sup>7</sup> and Johnsson et al.<sup>8</sup> used constant computer response times within each experimental condition, rather than a distribution of response times around some specified mean value. All this research can mislead readers who are unaware of the studies' limitations.

The present experiment examined user performance in a simple data entry task while the mean duration and the variability of computer response time was being manipulated. By carefully controlling these variables over a realistic range of stimulus values, some of the confusion caused by earlier experiments has been eliminated.

## II. METHODS AND APPARATUS

### 2.1 Task

Subjects were given a typed list of 1000 five-character letter groups to be used as input stimuli in the experiment.<sup>9</sup> All letter groups were printed in uppercase type. Half of the letter groups in the set were first-order approximations of English words, and the other half were zero-order approximations of English, based on the bigram frequency of the letters they contained. The zero-order and first-order letter

groups were randomly intermixed. Zero-order letter groups appear to be random sets of characters, and first-order letter groups appear to be very English-like,<sup>10</sup> though none of the groups used here was a real word. These stimuli were chosen to approximate the mixture of English words and code values entered by data entry clerks in large operations-support systems. The same set of input stimuli was used in every data collection session.

Subjects typed these letter groups one group at a time at a video display terminal (Hewlett-Packard 2621). They were instructed to work steadily at this task, and to complete as much of the work as possible within the allotted time. No specific incentives were offered to encourage them to work quickly. A single transaction proceeded as follows: The subject, on seeing the prompt character displayed on the terminal screen, typed the next letter group on the list, and followed it with a carriage return. This carriage return marked the beginning of the computer response time interval imposed by the experiment. When the appropriate amount of time had elapsed, the prompt character was again displayed at the beginning of the next line, signaling that the computer was ready to accept the next input. Until the prompt symbol was displayed, the keyboard was locked and it was impossible for subjects to type anything into the system.

Subjects could correct typing mistakes that they noticed before striking the carriage return with a "character erase" character "#", which erased the character immediately preceding it. It was not possible for subjects to cancel an entry after they struck the carriage return key, even if they noticed that the entry was incorrect. They were told to proceed to the next letter group if this occurred.

Subjects worked at each of the ten experimental conditions for 2 hours, or until they completed the list of 1000 letter groups, for a maximum total of 20 hours of data collection per subject. (Many subjects completed the list of letter groups in less than 2 hours in the shortest computer response time conditions.) All data collection sessions for individual subjects were separated by at least a full day, and only one subject took part in the experiment at a time. Each subject worked through the ten experimental conditions in a different, random order.

## **2.2 Computer response times**

Within each session, elements from a positively skewed distribution of response times were randomly presented to the subject. Distributions of real-world computer response time are usually strongly skewed in the positive direction. The shape of the distribution used here was approximately that of Chi-squared with 4 degrees of freedom, but its mean and standard deviation varied from condition to condition in

Table I—Computer response time distributions

| Condition | Range of $\bar{x}$<br>(in Seconds) | Range of<br>Standard<br>Deviation<br>(in Seconds) | Range of<br>Skew | Lowest Value<br>(in Seconds) | Highest Value<br>(in Seconds) |
|-----------|------------------------------------|---|------------------|------------------------------|-------------------------------|
| 2LO       | 2.002–2.021                        | 0.264–0.269                                       | 0.61–0.65        | 1.589–1.610                  | 2.720–2.734                   |
| 2HI       | 2.000–2.027                        | 0.793–0.803                                       | 0.59–0.68        | 0.786–0.803                  | 4.149–4.164                   |
| 4LO       | 4.006–4.020                        | 0.595–0.599                                       | 0.62–0.64        | 3.087–3.106                  | 5.611–5.616                   |
| 4HI       | 3.999–4.039                        | 1.788–1.800                                       | 0.62–0.65        | 1.283–1.292                  | 8.805–8.812                   |
| 8LO       | 8.009–8.032                        | 1.329–1.347                                       | 0.61–0.63        | 5.966–5.992                  | 11.568–11.583                 |
| 8HI       | 8.020–8.105                        | 4.016–4.072                                       | 0.60–0.62        | 1.912–1.926                  | 18.700–18.716                 |
| 16LO      | 15.992–16.035                      | 2.972–3.079                                       | 0.58–0.63        | 11.466–11.481                | 23.979–23.994                 |
| 16HI      | 16.073–16.236                      | 9.046–9.128                                       | 0.58–0.63        | 2.415–2.419                  | 39.928–39.929                 |
| 32LO      | 32.059–32.221                      | 6.720–6.881                                       | 0.60–0.64        | 21.874–21.883                | 49.806–49.829                 |
| 32HI      | 31.792–32.409                      | 20.180–20.704                                     | 0.59–0.68        | 1.619–1.630                  | 85.416–85.437                 |

Note: In the column labeled "Condition", the integer indicates the nominal mean response time and the "LO" and "HI" suffixes indicate the "low" and "high" variability conditions, respectively.

the experiment. Five different mean computer response times were used: 2, 4, 8, 16, and 32 seconds. Two different variability conditions were imposed on each mean value, for a total of 10 different experimental conditions. Table I summarizes the stimuli used in the experiment. Stimulus values varied slightly from subject to subject because of differences in the number of transactions completed in each session across subjects. However, the ranges of values shown in Table I encompass the data for all subjects. The range of mean response times used here was set after meeting with many designers of large operations systems. Though the response time design goals for such systems are normally set at about 4 seconds, the range used here is an accurate reflection of the values encountered by developers during system performance tests.

Response time variability was calibrated in psychological units across the different conditions. In the "low" variability condition, one standard deviation of the computer response time distribution was set to equal 1.0 jnd (just noticeable difference) of computer response time at each of the mean values used. In the "high" variability condition, the standard deviation of each computer response time distribution was set to equal 3.0 jnd at each mean value. This equated the perceived variability within each low and high condition across the different mean values. The ranges of computer response time within each condition resulting from this scaling are shown in Table I. Just noticeable differences of computer response time were calculated using the measurements of Butler et al.<sup>11</sup>

### 2.3 Computer system

A PDP-8\*/E laboratory computer controlled the subject's terminal.

\* Trademark of Digital Equipment Corporation.

It timed the prescribed computer response time intervals, recorded these values, and also recorded the subject's response time after the prompt character was displayed, the subject's typing time, and all characters typed during each transaction. These data were sent to a host computer and filed there as they accumulated.

## **2.4 Subjects**

The subjects in this experiment were six experienced word processing clerks employed by AT&T Bell Laboratories in Piscataway, New Jersey. All were proficient in computerized text preparation, and all had previously used a video display terminal of the same model as the one used in this study (Hewlett-Packard 2621).

## **III. RESULTS**

Three different measures of user performance were recorded during each transaction: user response time, user typing time, and user input errors. For the purposes of this study, user response time is defined as the time between the display of the prompt character on the subject's terminal and the typing of the first character entered by the subject. User typing time is the interval between the first character typed in each transaction and the typing of the carriage return that marks the end of the transaction. User errors were compiled here by calculating the percentage of incorrect entries made by each subject in a given session. Entries were judged incorrect if they were misspelled, not entered in the prescribed sequence, or omitted. Results for each of the three measures are presented below.

Neither mean input error rate nor mean typing time per transaction was significantly affected by the computer response time variables studied here. Error rates were quite variable within each condition, but mean error rate remained approximately constant at about 2 to 4 percent for mean computer response times ranging from 2 to 32 seconds in both the high- and the low-variability conditions. Typing time was also approximately constant at about 1.25 s/transaction across all experimental conditions.

The plots of mean user response time as a function of mean computer response time shown in Fig. 1 are somewhat more interesting. While the variability in these data is quite large, a regular relationship between these two variables appears to be present. User response time, which is about 1.0 to 1.25 seconds at a mean computer response time of 2 seconds, increases steadily to about 3.5 to 4.0 seconds at a mean computer response time of 32 seconds. Also, the high-variability condition regularly incurs user response times about 0.75 second longer than those resulting during the low-variability condition.

Logarithmic curves are one way to conveniently represent the data in the two variability conditions:

$$\text{low variability: } Y = 2.179 \log X + 0.075 \quad r^2 = 0.913$$

$$\text{high variability: } Y = 2.518 \log X + 0.465 \quad r^2 = 0.990,$$

where  $Y$  equals mean user response time and  $X$  equals mean computer response time.

These equations were calculated by transforming the computer response time variable to a logarithmic scale, and then performing a linear least-squares regression on the transformed data. The resulting plot, including the regression lines, is shown in Fig. 2.

A repeated-measures analysis of variance showed that the effect of computer response time variability was significant ( $p < 0.05$ ) and that the effect of mean computer response time was marginally significant ( $p = 0.053$ ). Both of these variables accounted for a fairly small percentage of the total variance in the experimental data, however. The data were noisy, and between-subjects variability accounted for about 84 percent of the total variance.

#### IV. DISCUSSION

In general, the data on errors and typing time gathered in the present study concur with the results of earlier studies using similar tasks.<sup>12</sup> Because these measures are not affected by different computer response time conditions, they are of little interest here.

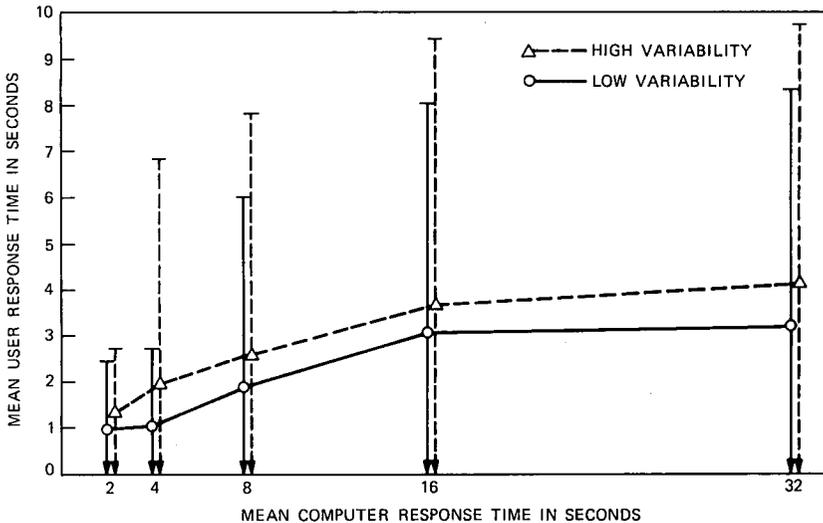


Fig. 1—The average user response time plotted as a function of mean computer response time. The error bars indicate  $\pm 1$  standard error.

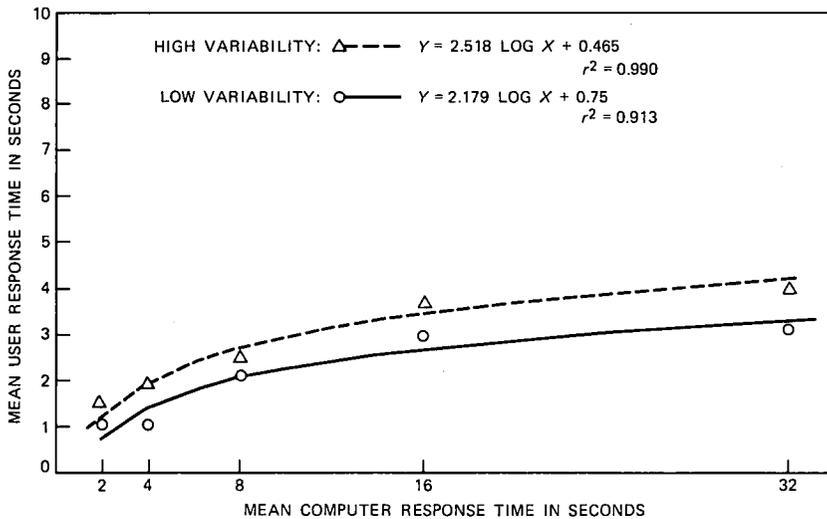


Fig. 2—The average user response time plotted as a function of mean computer response time. The lines drawn through the data points are linear-log least-squares regression lines.

The effect of different mean computer response times on user response times is another, more interesting matter. Several studies have made specific conclusions about the relationship between these two variables. Some of these conclusions are affirmed by the present study, and some are not.

One striking finding of several papers has been that user performance is optimal at computer response times of intermediate duration,<sup>13-15</sup> usually the values have been found to be in the 4- to 5-second range. There is no evidence of any decrement in performance at short computer response times in the data from the present experiment. Rather, performance (as measured by the user's response time) gradually worsens over the range of computer response times used. (This measure of user performance worsens over the range of average computer response times from 4 to 32 seconds. Changes in user performance over the 2- to 4-second range of stimuli are less obvious.) The important difference between the present study and these earlier experiments is that all the older papers reporting this finding were investigating specific problem-solving tasks, while the present study used a simple data-entry task. It could be that the subjects in these problem-solving studies required some amount of preparation or "think time" between transactions; they were planning their next command. With the data-entry task used in the present study, subjects needed no think time. They knew what their next command had to be and merely waited for the opportunity to enter it. The difference found

here provides confirmation for Miller's speculation (see Ref. 16) that different commands have different response time requirements for users, depending on the cognitive load that they impose. It also points out the importance of considering the task type of a given command before extrapolating experimental data from other types of tasks to it.

Another widely cited set of data relating computer response time to user response time originated in a study by Boies and Gould,<sup>5</sup> and has also been reported by Boies<sup>17</sup> and Doherty and Kelisky.<sup>18</sup> This was a survey-type study in which all usage was monitored in a general-purpose research computer system. Doherty and Kelisky found that the expected user response time was equal to 15 seconds plus whatever the computer response time was for that transaction.<sup>18</sup> The results of the present study are clearly different from this, as shown in Figs. 1 and 2. This difference is not surprising, since Doherty and Kelisky described responses during a heterogeneous mixture of user command types, and because they correlated user think times with the single preceding computer response time, rather than with the average response time of the system.

An important finding of the present study is that users appear to respond to the computer's average response time, rather than to the last single response time experienced. One extreme hypothesis is that a user's response time is determined by the computer response time that immediately precedes it, and its opposite extreme is that the user's expectation of what the computer's response time should be, based on his or her total experience with the system, determines what his or her response will be. The first of these hypotheses is implicit when one uses survey data like that cited by Doherty and Kelisky.<sup>18</sup> One must assume that the mechanism mediating the user's response has no memory, and that each transaction is independent of all others that have preceded it.

A straightforward test of the two alternatives can be made using data from the present study. To make this test, all individual transactions in the high-variability condition with computer response times of 3.5 to 4.5 seconds were isolated, and the average user response time of each individual subject for these transactions was compared across mean computer response time conditions. The results of the test were qualitatively similar for all subjects. Data from two of them—the subject with the largest performance variation across all conditions, and the subject with the smallest overall performance variation—are shown in Figs. 3 and 4, respectively. These figures are plots of mean user response time as a function of mean computer response time. It is clear from these data that a user's average response time to these isolated transactions of 3.5- to 4.5-seconds duration is not constant across conditions, and that the important determinant of the user's

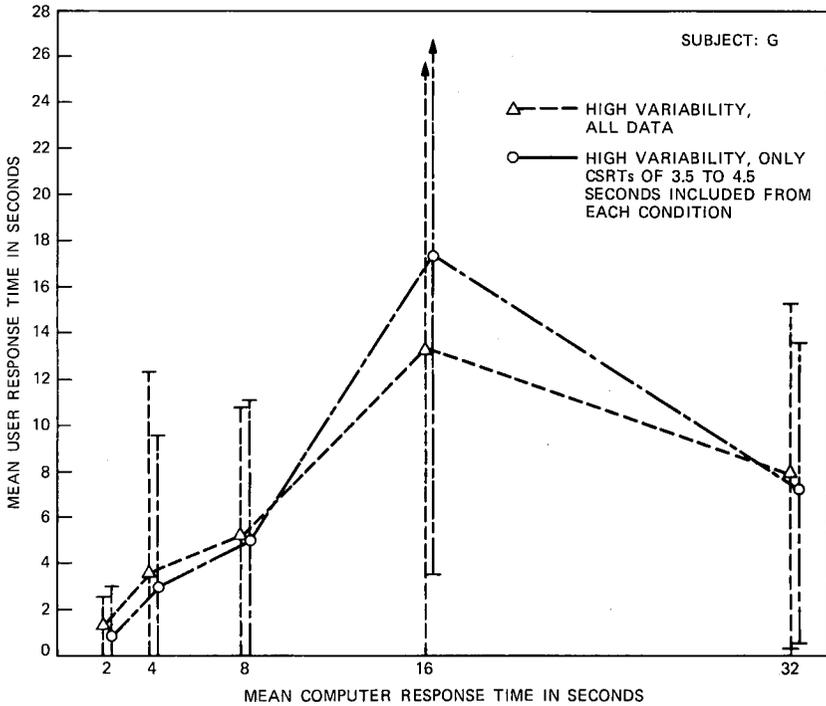


Fig. 3—Average user response time plotted as a function of average computer response time for the subject G.

response time here is the *mean* computer response time within each session.

#### 4.1 Computer response time variability

Uniformly increasing the variability (in psychological units) of the computer response time distributions that the subjects experienced succeeded in uniformly increasing mean user response times by about 0.75 second across all conditions. Unfortunately, it is not possible to quantify this relationship with only two levels of variability in the study. This point is probably worth further examination in the future.

#### 4.2 Conclusions

Systematically manipulating the response time characteristics of a computer system while users performed a simple data-entry task has shown that increases in the response time of the system cause a slow, gradual, and not very large degradation in user performance, as indicated by the user's response time. Also, increasing the variability of computer response time significantly degraded user performance, ac-

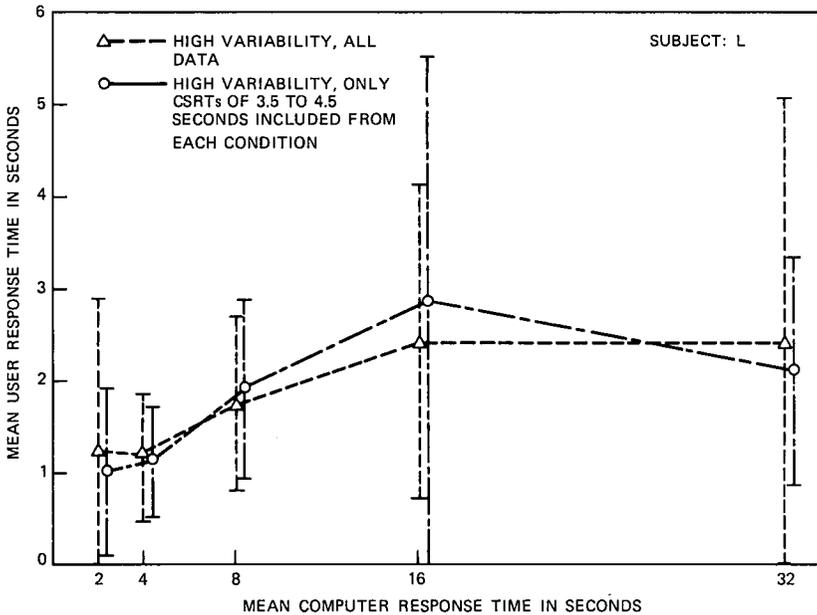


Fig. 4—Average user response time plotted as a function of average computer response time for the subject L.

ording to the same measure. There is no reason to believe that these results can be applied directly to user tasks requiring more cognitive effort, or to tasks in which the user's response depends on the content of the preceding computer output. These results should, though, set a baseline to which user performance with other, presumably more complex, user tasks can be compared.

The psychological basis for the effects seen here is not obvious. At first glance, it could be seen as only a problem of attention for the subjects; the long computer response time intervals could simply be making their responses more lethargic. Two facts, though, suggest that this simple hypothesis is not adequate: First, if attention was a problem, input error rates would probably rise along with each user's response time. This did not happen in the present experiment. Second, the high level of variability seen in these data would probably not be expected if something as simple as attentiveness were mediating subjects' responses. The subjects in the experiment were a fairly homogeneous population of experienced computer users, yet the variability of response time between users and even in each user's case was high. This suggests that some more complex factor or factors are important to the subjects' responses. Further research is needed before we will know the basis for the results of this study.

The most striking finding of this study is how little user performance at this task is influenced by computer response time. Of all the measures of user performance tested, only one was affected at all by computer response time, and the performance degradation shown by this measure as computer response time increased was quite small. The second important finding of the present study is that computer users' performance appears to be related to *average* computer response time, rather than to the duration of the immediately preceding computer response time interval. This finding has important methodological implications for future studies.

## REFERENCES

1. J. C. R. Licklider, "Man-Computer Symbiosis," *IEEE Trans. Hum. Factors Electron., HFE-1* (March 1960), pp. 4-11.
2. J. C. R. Licklider, "Man-Computer Partnership," *Int. Sci. Technol.*, 33 (May 1965), pp. 18-26.
3. J. R. Carbonell, J. I. Elkind, and R. S. Nickerson, "On the Psychological Importance of Time in a Time Sharing System," *Hum. Factors*, 10 (April 1968), pp. 135-42.
4. S. E. Engel and R. E. Granda, "Guidelines for Man/Display Interfaces," IBM Tech. Report TR00.2720, December 19, 1975.
5. S. J. Boies and J. D. Gould, "User Performance in an Interactive Computer System," Proc. 5th Annual Princeton Conf. on Info. Sci. and Systems, Princeton, NJ, March 25-26, 1971.
6. R. E. Barber, "Response Time, Operator Productivity, and Job Satisfaction," Ph.D. dissertation, NYU Graduate School of Business Administration, 1979.
7. M. C. Grignetti and D. C. Miller, "Modifying Computer Response Characteristics to Influence Command Choice," Proc. Conf. on Man-Computer Interactions, Cambridge, UK, September, 1970, pp. 201-5.
8. B. Johnsson, B. Andersson, and S. Wallin, "Man-Computer Communication Through Alpha Numeric Display Terminals," Int. Conf. on Cybernetics and Society, Tokyo-Kyoto, Japan, 1978, 2-3: pp. 1278-83.
9. K. Hirata and M. P. Bryden, "Tables of Letter Sequences Varying in Order of Approximation to English," *Psychonom. Sci.*, 25, No. 6 (December 1971), pp. 322-4.
10. C. N. Cofer, "Properties of Verbal Materials and Verbal Learning," *Woodworth and Schlossberg's Experimental Psychology*, 3rd ed., ed. J. W. Kling and L. A. Riggs, New York: Holt, Rinehart, and Winston, 1971, pp. 896-8.
11. K. A. Butler, G. L. Felfoldy, S. E. Simms, and J. S. Swenson, unpublished work.
12. J. D. Williams, J. S. Swenson, J. A. Hegarty, and T. S. Tullis, unpublished work.
13. B. W. Boehm, M. V. Seven, and R. A. Watson, "Interactive Problem-Solving: An Experimental Study of 'Lockout' Effects," AFIPS Conf. Proc., Atlantic City, NJ, May 18-20, 1971, 38: pp. 205-10.
14. M. A. Morfield, R. A. Wiesen, and M. Grossberg, "Initial Experiments on the Effects of System Delay on On-line Problem-solving," Lincoln Laboratory, MIT, Lexington, MA, Tech. Report No. ESD-TR-69-158, 1969.
15. M. Grossberg, R. A. Wiesen, and D. B. Yntema, "An Experiment on Problem Solving With Delayed Computer Responses," *IEEE Trans. Syst. Man Cybern., SMC-6* (March 1976), pp. 219-22.
16. R. B. Miller, "Response Time in Man-Computer Conversational Transactions," AFIPS Conf. Proc., San Francisco, CA, December 9-11, 1968, 33, pp. 267-77.
17. S. J. Boies, "User Behavior on an Interactive Computer System," *IBM Syst. J.*, 13, No. 1 (1974), pp. 2-18.
18. W. J. Doherty and R. P. Kelisky, "Managing VM/CMS Systems for User Effectiveness," *IBM Syst. J.*, 18, No. 1 (1979), pp. 143-63.

## AUTHOR

Thomas W. Butler, B.Sc. (Psychology), 1972, Michigan State University;

Ph.D. (Psychology), 1977, Brown University; University of California at Berkeley, 1976–1978; AT&T Bell Laboratories, 1978—. At the University of California Mr. Butler was a member of the Physiology–Anatomy Department. Since joining AT&T Bell Laboratories, he has worked on methods for the human estimation of work time, and on the effects of computer response time on user performance. He is currently a member of the Human Performance Engineering department, where his work includes a variety of projects related to the *UNIX*<sup>™</sup> operating system.