WADC TECHNICAL REPORT 59-7

THE DIFFERENTIAL EFFECTS OF SPEED AND LOAD STRESS ON TASK PERFORMANCE

Gabriel Jeantheau

Laboratory of Aviation Psychology

The Ohio State University
and
The OSU Research Foundation

JULY 1959

Contract No. AF 33(616)-3612
Project No. 7184
Task 71583

AERO MEDICAL LABORATORY
WRIGHT AIR DEVELOPMENT CENTER
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

1,300 — February 1960 — 23-620

FOREWORD

This report covers part of the research on man-machine systems being conducted by the Laboratory of Aviation Psychology and the Department of Electrical Engineering of The Ohio State University, with Dr. George E. Briggs as Principal Investigator. The objectives of this research are: (1) the development of new human factors methodology for studying man-machine systems, (2) the application of new methodology to several different types of systems in order to modify and improve the validity and generality of concepts, (3) the development of human factors principles for the analysis and synthesis of systems, and (4) the formulation of human factors principles and information in terms compatible with standard engineering practice.

The present report was prepared for the Engineering Psychology Branch, Aero Medical Laboratory, Directorate of Laboratories, Wright Air Development Center, under Contract No. AF 33(616)-3612, Project 7184, Task 71583, with Dr. James C. McGuire acting as Task Scientist. This work was initiated under Contract No. AF 33(616)-43 with Dr. Ralph W. Queal, Jr. acting as Project Scientist and Dr. Paul M.Fitts as Principal Investigator.

The author expresses his appreciation to Professor Paul M.Fitts, his advisor, for his many helpful suggestions and encouragement.

This report is based in part on a thesis submitted to the Graduate School of The Ohio State University in partial fulfillment of the requirements for the degree Master of Arts.

The author is now with Dunlap and Associates, Inc., Stamford, Connecticut.

ABSTRACT

This study investigated effects of speed-load stress (input rate and variety) on performance in an information-processing task requiring discrimination of pairs of identical symbols within matrices of three levels of word length and five levels of presentation rate. Error scores and a derived information measure were used to assess performance.

Results indicated that increases in both speed and load seriously impair performance. The data fitted to an information specification of the stimulus situation failed to provide a unitary function. Information scores plotted against mean number of paired comparisons per second, however, yielded a single function which described all conditions and accounted for the high interaction component.

It was concluded that (a) specification of the stimulus situation alone does not adequately explain performance, and (b) use of information measures in tasks of this type adds little or nothing to such explanations. A more pertinent consideration is the actual operations required in responding to the task situation.

PUBLICATION

This report has been reviewed and is approved.

FOR THE COMMANDER:

Walter F. Grether

WALTER F. GRETHER
Director of Operations
Aero Medical Laboratory

TABLE OF CONTENTS

THE DIFFERENTIAL EFFECTS OF SPEED AND LOAD STRESS ON TASK PERFORMANCE

INTRODUCTION

Technological advances of the past decade have enforced upon psychologists a renewed interest in performance on complex tasks. Traditional considerations of this problem followed two main directions. One emphasis was on the analysis of the effects on performance of a concurrent second task, the other on studies of individual differences in ability to perform two tasks at once. Neither of these approaches proved very fruitful, and no general concepts emerged which were able to stimulate further theoretical or experimental efforts. Not until World War II, with the development of complex man-machine systems, did this problem come into the foreground again.

Many present-day man-machine systems incorporate human operator tasks which require the individual to handle several information sources simultaneously. Such tasks place high demands on the operator in terms of the required rate and variety of responses. Consider the task of the radar air traffic controller. The controller must accept each aircraft entering the control zone and guide its flight path to a successful landing. A number of successive decisions must be made regarding each of these aircraft, based upon both the flight characteristics of the individual plane and its status with reference to the ground and other aircraft in the system. The latter consideration involves the coordination of many individual decisions into a single acceptable pattern in order to maintain an even flow of aircraft through the system. Thus, the task is a dynamic one, not of the static problem-solving or puzzle-box variety. The operator is part of an ongoing system and he must direct its flow. Although his activity is a directive one, the system is externally paced and events occur at a rate over which he has no control. He must, therefore, adapt to the demands of the system as well as exercise control over it.

In order to study behavior in tasks of this type, it is essential that we be able to specify the factors that determine the complexity of various aspects of the task such as its perceptual or input components. A number of factors have been studied in the past which were thought to contribute to the complexity of this type of task. Among these are the physical intensity of the signals, figure-ground relationships within the visual field, and spatial and temporal relations between subject and field. A variable more pertinent to tasks of a dynamic character is one which describes the rates at which display changes occur and the rates at which responses must be made. This variable, which we can refer to as the "speed" characteristic of the task, can be stated more simply as the number of signals occurring per unit time. A more complete description of the task requires the definition of a second variable. Conrad has called this variable "load" and defined it as "the number of separate streams of signals that occur independently of each other but require simultaneous consideration" (1951). The speed-load concept has been the subject of a series of investigations by Mackworth and Conrad (Conrad, 1951; Conrad, 1955; Mackworth & Mackworth, 1956; Mackworth & Mackworth, 1957). Their studies attempt to determine the effect of speed and load on perceptual complexity.

Manuscript released by the authors 15 December 1958 for publication as a WADC Technical Report.

The present study is a modification of the type of task used by Mackworth. In the latter task, S made an individual response to each separate source. Many tasks, however, are such that information from several sources must be combined and a summary statement made about the total complex of information.

For the purpose of describing this extension of the traditional speed-load task, several concepts drawn from information theory will be developed. "Alphabet," as used in information theory, refers to any set of coding symbols which can be recognized as a set and given a name. Common examples of these are the English letter alphabet, composed of 26 symbols, and the Arabic numeral alphabet, comprising 10 symbols. Binary alphabets contain only two symbols. The term "word" refers to a group of symbols drawn from any given alphabet. Words can be defined by their length or the number of symbol positions involved in their make-up. As an illustration, a teletype tape has five positions in which a hole can be punched or not punched. Any grouping of pulses in these five positions can be considered a word. In the type of task used by Mackworth, the term "word" can be used to identify a group of six symbols displayed on each stimulus card in a window. Aggregates of words are referred to as "messages." Defined in this way, each window in a Mackworth display contains a message consisting of two six-symbol words. In this sense, S's task in any information-processing task is to deal with messages which may be further broken down into component words and symbols. This framework of terminology provides a means of categorizing the information content of the display in a speed-load task. Speed can be specified per symbol, per word, or per message.

Finally, in summary, the present study was planned to investigate the independent effects of variations in task speed and task load in an information-processing task. Load was varied in terms of word length with alphabet size and words per message held constant. Speed was varied as the number of words per minute.

METHOD

Apparatus

The apparatus consisted of an Esterline-Angus chart drive modified to include a flat, vertical viewing surface. The flat part of the drive was screened with an aluminum shield with a 4X6-in. Plexiglas window inset. The window was ruled vertically into four columns of equal width, such that symbols moving past the window appeared as four "separate streams of information."

The device was driven by a variable-speed tape puller which permitted rate settings over a continuous range from 0 in./sec. to well beyond the speed requirements of the study. Four sets of stimuli, or alphabets, of code symbols were used: the S alphabet (variations of squares), the N alphabet (numerals), the C alphabet (variations of circles), and the L alphabet (letters). These alphabets are shown in Fig. 1. Each alphabet contained four alternative symbols; therefore each code symbol provided 2 bits of information. The symbols were printed on rolls of Esterline-Angus chart paper in messages consisting of 8, 12, or 16 symbols (see Fig. 2). Each message consisted of four words (rows), and each word consisted of either 2, 3, or 4 symbols (columns). A roll contained 85 such messages, each separated from the next by a horizontal red line, with a fixed interval (four blank word spaces) between messages.

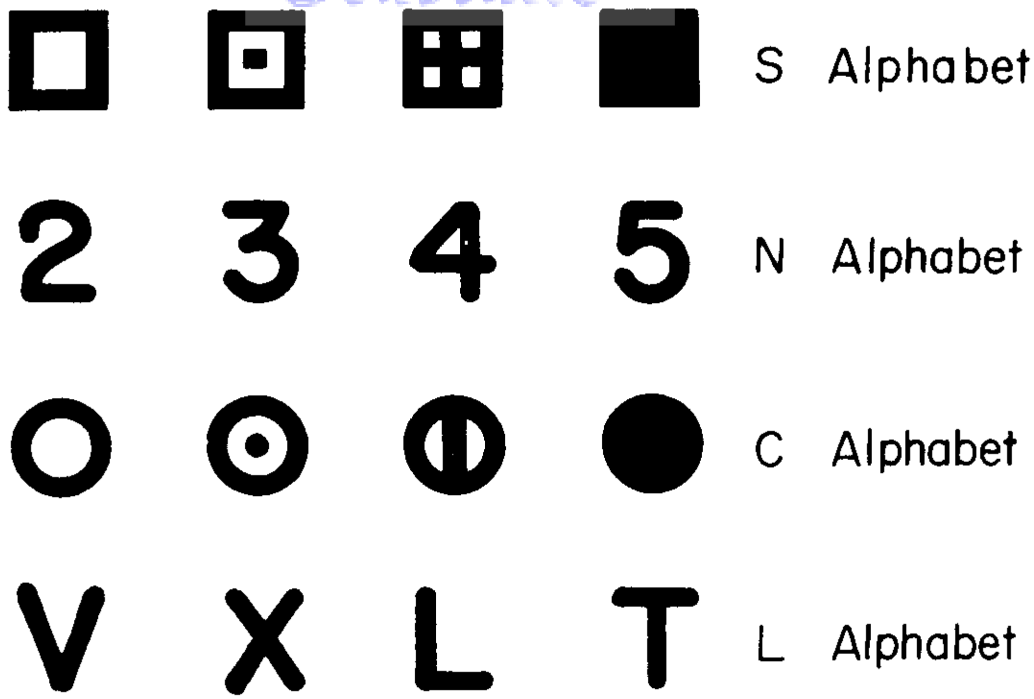

Fig. 1. Alphabets used in word construction.

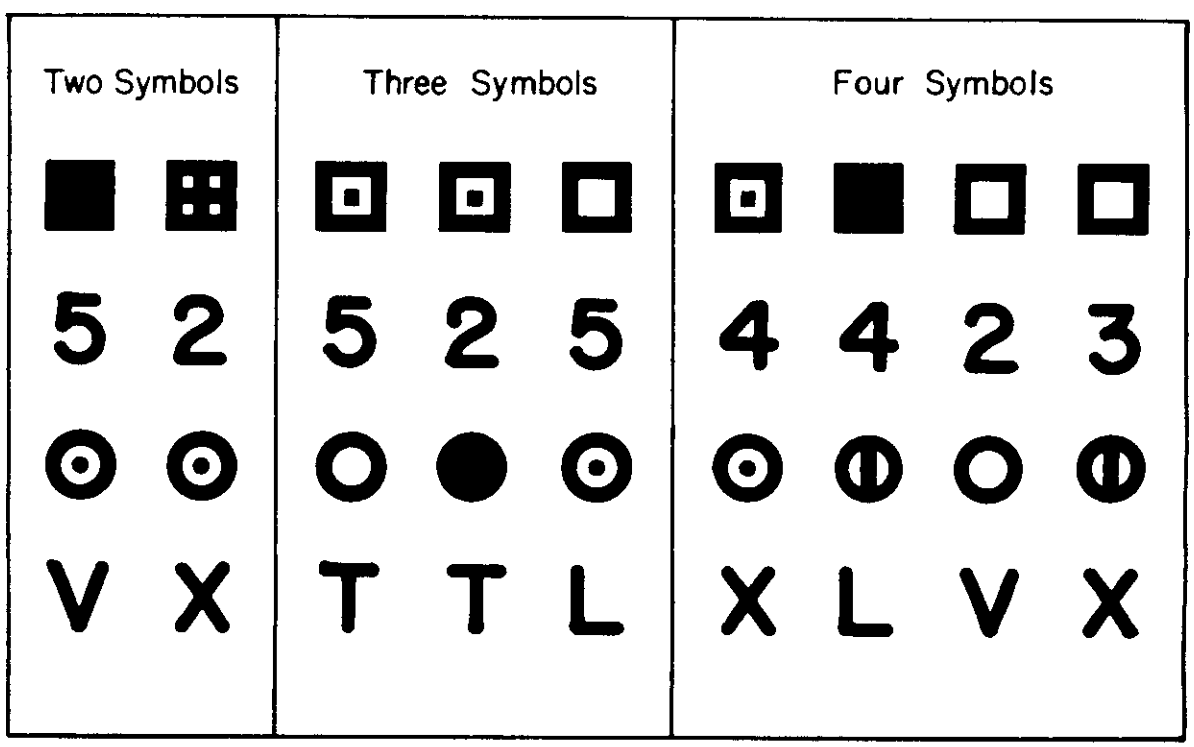

Fig. 2. Message samples.

Each message consisted of four words, but each word was written with a different one of the four alphabets, namely, an S-word, an N-word, a C-word, and an L-word arranged one above the other in that order. Words were written in two ways, either (a) with no duplicate symbols or (b) with a single pair of duplicate symbols, i.e., no word contained a quadruplet, a triplet, two pairs, or other forms of duplicate symbols.

Pairs of duplicate symbols were randomized, using a table of random numbers, with respect to the following:

a. Number of words containing a duplicate pair of symbols per message
b. Word in which pairs occurred
c. Symbol within the word which was used to form the pair
d. Symbol positions (columns) within which members of the pair occurred

If the symbols making up words had been selected on a truly random basis, the frequencies of different numbers of pairs (from 0 to 4) would have been unequal. Therefore, in order to avoid any discrimination biases that might have resulted from such a distribution, the frequencies of pairs per message were made equal on each roll. This restriction, it should be noted, results in a slight reduction in the information content of the messages.

Subjects

The Ss were 10 students at The Ohio State University. These students were part-time employees of the Laboratory of Aviation Psychology and had excellent orientation to research methodology. Although they were naive with respect to the present task, they had been trained to respond accurately and rapidly in other complex situations.

Procedure

Experimental conditions included five rates, 17, 20, 24, 30, and 40 messages per minute. These rates correspond to exposure times of 3.5, 3.0, 2.5, 2.0, and 1.5 sec. per message. Three levels of complexity or lengths of words were investigated. These included 2, 3, and 4 symbols per word (i.e., 8, 12, or 16 symbols per message). The five levels of rate and three levels of load were combined factorially and each S served under all 15 conditions. The data as presented in Tables 1 and 2 represent this design graphically. The Ss were given 15 training trials, one under each condition, and similarly, 15 experimental trials. The 30 trials were run in four sessions of eight and seven trials alternately. A 2-min. rest interval occurred between trials. The order of presentation of conditions was random and different for each S, and a different order was used for training and experimental trials.

Each S was given the following instructions prior to the first training trial:

> In this experiment we are trying to find out something about how people process information. On the device you see in front of you, groups of symbols will move downward in the window. Each group consists of four rows and either 2, 3, or 4 columns. You are to think of these groups as messages, where each row represents a

word. (S is shown samples of messages, one of each word length.) The messages are separated by horizontal red lines. You will notice that in some words (rows) there are pairs of identical symbols. These pairs can either be adjacent or split, but the symbols will be identical. There will never be more than one pair in any one word. It is possible that an entire message may contain no pairs, i.e., every symbol in every word will be different.

Now, consider yourself in a message center, receiving streams of coded information in 2, 3, or 4 letter words in four-word messages. The thing you are interested in is the number of words containing pairs of identical symbols. The object of this task, then, is to count the number of symbol pairs in each message as it appears in the window. As soon as you know how many there are, call out the total for that message. There are only five possible answers: 0, 1, 2, 3, or 4. There is equal likelihood that any number of pairs can occur, including zero.

We are going to present the messages at several different rates. In some cases they will go by rather quickly. It is important that you try to respond to every message, even if you must guess. Are there any questions?

The S's response to each of 75 messages was recorded on each trial. Omissions were scored as errors. Actually, during the course of the experiment, omissions occurred only infrequently during training trials and never occurred during the experimental trials.

RESULTS

Per cent error scores, pooled over Ss, are presented in Fig. 3. It is evident that as rate increased, per cent error increased. There was a marked difference, however, between the high and low load conditions. On the one hand, in the two-symbol word length condition, performance was never seriously affected over the range of rates used in this study. Even at the highest rate, 40 messages per minute, errors did not reach 10%. On the other hand, performance suffered extensively at high rates in the three- and four-symbol word length conditions. Errors exceeded 50% in the four-symbol condition at the 40 message per minute rate.

Error scores on each trial for each S were converted to an information measure. The computational method used was to tally, in a 5X5 stimulus-response matrix, the actual number of duplications against the reported number of duplications per message. The mean of the $p \log p$'s across rows (average response equivocation) was subtracted from the total $p \log p$'s in the columns (relevant stimulus information). The difference thus obtained can be considered relevant information transmitted in terms of the five possible and equally likely alternative responses. With five alternative responses, the maximum score that could be achieved in an errorless trial was 2.322 bits transmitted. Graphically, the data as presented in Fig. 4 take much the same form as per cent errors. Again the high load conditions show serious decrement with increases in rate.

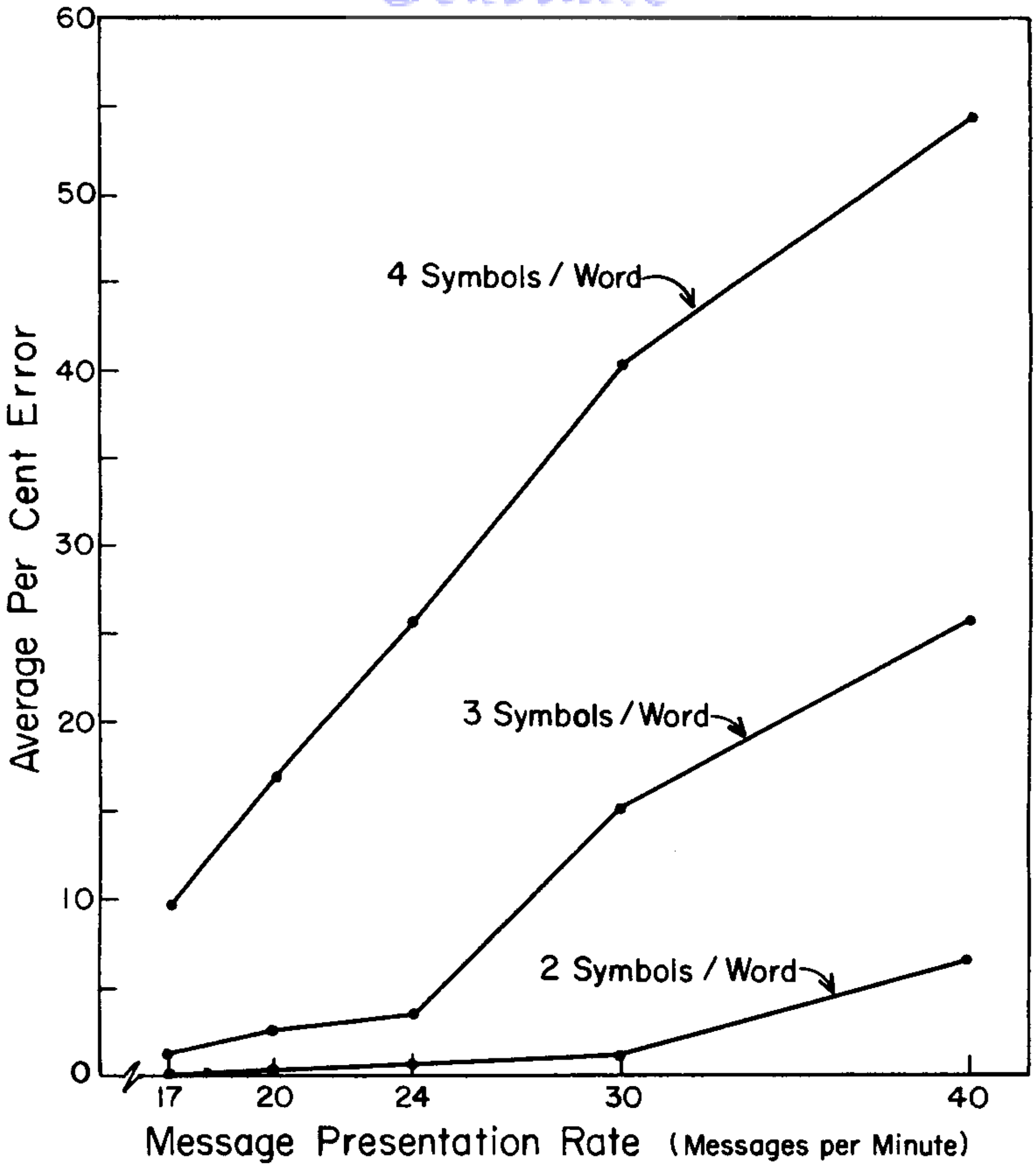

Fig. 3. Per cent error as a function of speed.

Bartlett's test for homogeneity of variance performed on these information measures indicated a strong heterogeneity of variance. Since one cell contained zero variance, the test was run on the remaining 14 conditions and yielded a χ^2 of 94.45 ($\underline{P} < .001$).

This value precluded the use of parametric analysis of variance techniques in further analysis of the data. A nonparametric test, Friedman's χ^2_r, was

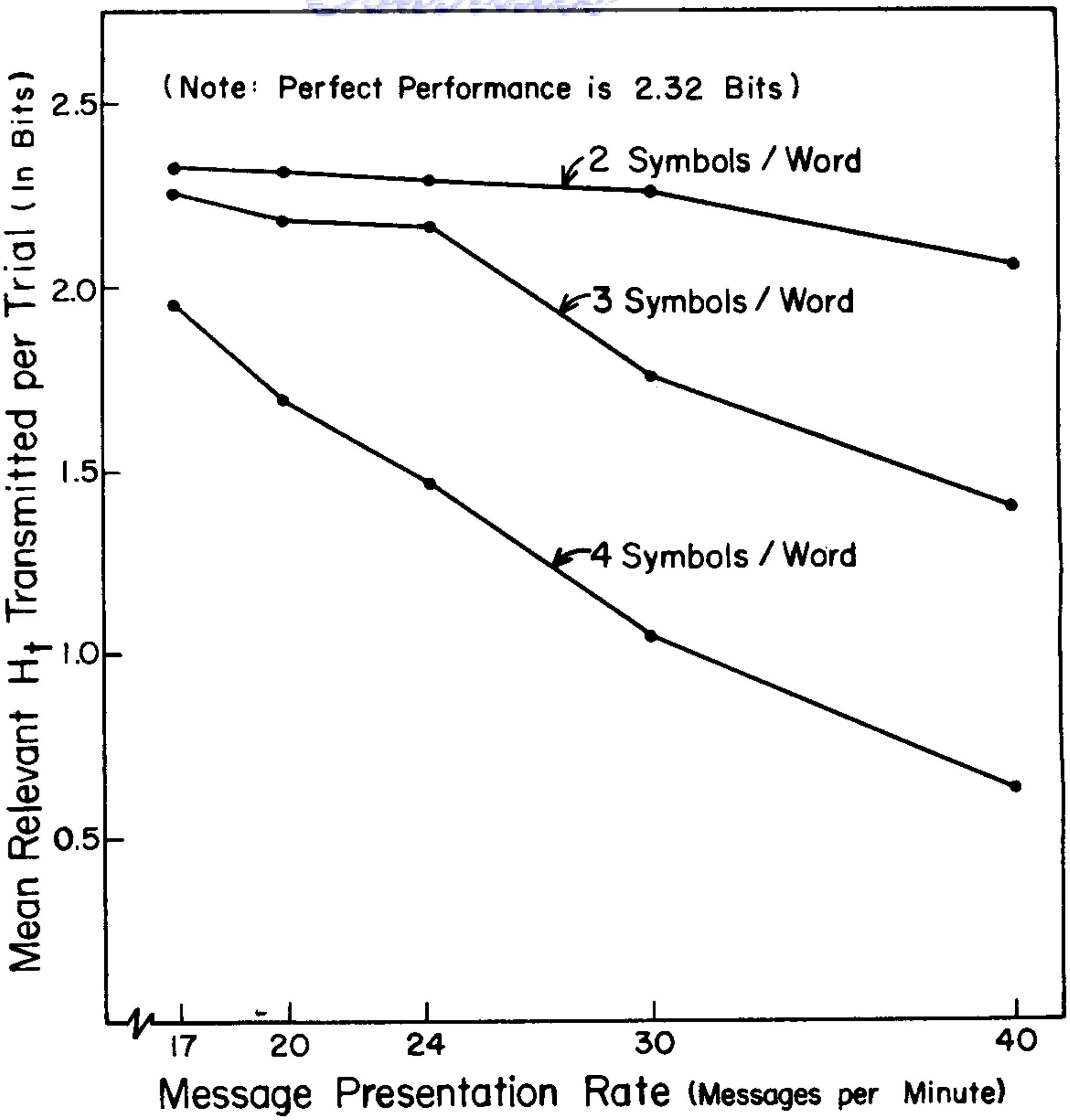

Fig. 4. Relevant information transmitted per trial as a function of speed for three levels of word length.

employed to test differences between conditions. Table 1 shows the results of this analysis. All values are significant at $\underline{P} < .001$ except the slowest rate and the lowest load. Interaction effects were tested by using Wilcoxon's extension of the Friedman method. This test indicated a highly significant interaction ($\underline{P} < .001$) of the speed and load variables.

Table 1

Summary Comparison of Conditions Using Bits of Information Transmitted Per Trial as the Criterion

Conditions Compared	Level of Orthogonal Variable	χ_r^2 [a]	P
Rate of Message Presentation	All Levels Pooled	26.68	< .001
	2 symbols/word	11.78	< .02
	3 symbols/word	31.06	< .001
	4 symbols/word	36.56	< .001
Number of Symbols Per Word	All Levels Pooled	19.20	< .001
	17 messages/min.	7.20	< .05
	20 messages/min.	16.80	< .001
	24 messages/min.	16.25	< .001
	30 messages/min.	19.05	< .001
	40 messages/min.	20.00	< .001
Interaction	----	47.88	< .001

[a] The Friedman Two-Way Analysis of Variance Test was employed to evaluate the main effects: the Wilcoxon extension of this same test was used to evaluate the interaction (Siegel, 1956).

Error frequencies were tabulated as a function of the number of pairs per message. These are shown in Table 2. In the two-symbol word condition, messages with four pairs were responded to almost as accurately as those with only one pair. In all three load conditions, messages containing three pairs were missed most frequently.

Table 2

Error Frequencies as a Function of Number of Actual Pairs Per Message

Symbols Per Word	Pairs 0	1	2	3	4
2	1	8	22	26	9
3	37	72	73	96	83
4	72	151	257	334	270

DISCUSSION

The data as presented in Fig. 3 agree generally with the results of previous studies in this area. At given rates of presentation, increases in load result in large decrements in performance; conversely, with load held constant, rate increases also adversely affect performance. These effects appear to be compounded, i.e., performance as measured by per cent errors falls off very rapidly indeed when both speed and load are increased. In this connection it should be noted that the test for speed x load interaction effects yielded the largest single component of the variance (see Table 1). Such effects are not unexpected. An increase in the number of sources of information in effect reduces the length of time available to process a single source, resulting in errors comparable to those due to an increase in speed alone.

The frequency of errors in responding to identical symbols is unusual and somewhat unexpected. Two trends appear in these results (see Table 2). One effect is that accuracy is greatest for zero pairs; Ss apparently do not report identical pairs when there are none. It may be concluded that errors result from not seeing pairs when they are present (not from seeing pairs that are not present). The other effect is that four identical pairs are reported more accurately than three. In this case the most likely explanation appears to be faster scanning of the message when many pairs are present; if S finds a pair of identical symbols in a word, he can immediately go on to the next word (since there is never more than one pair per word), thus affecting a savings in scanning.

The concept of adjacency may explain some of the effects already noted. Since the occurrence of pairs was randomized in the construction of stimulus words, there were fairly equal frequencies of adjacent pairs and split pairs from word to word. Although an analysis of errors from this viewpoint was not possible, it is not unreasonable to assume that adjacent pairs are more easily and quickly seen than split pairs. This could account for the relatively high performance in the two-symbol condition. In the two-symbol condition a split pair cannot occur. In the three-symbol condition, however, the probability of a split pair is 1/3 and in the four-symbol condition this probability increases to 1/2. Thus, adjacency may account for a large part of the difference between conditions.

The development of the word-message concept earlier in this report was made to permit clearer specification of the information in the stimulus situation. The difficulty of a dynamic task is not only a function of speed and load, but also of the information content of the display. Displayed information can be specified in a number of ways, namely, bits per symbol, bits per word, or bits per message. It was felt that this might provide some useful prediction device, particularly a means of simplifying the complex set of curves in Fig. 3 and 4. A plot was made of performance as a function of the number of bits per second for the various speed-load conditions. The result is shown in Fig. 5. It will be seen that the data now come close to being fitted by a single function in which, after a speed-load combination of about 10 bits/sec. is reached, errors increase almost linearly with increased information input.

A somewhat different approach was taken by Sidorsky (1954) in a study in which Ss were also required to discriminate and report identities among groups of coded symbols. He specified task difficulty in terms of the number of possible pairs of stimuli to be compared, hypothesizing that the best prediction in this type of situation was a function of the number of paired comparisons S was required to make in order to respond correctly. Theoretically, this figure is

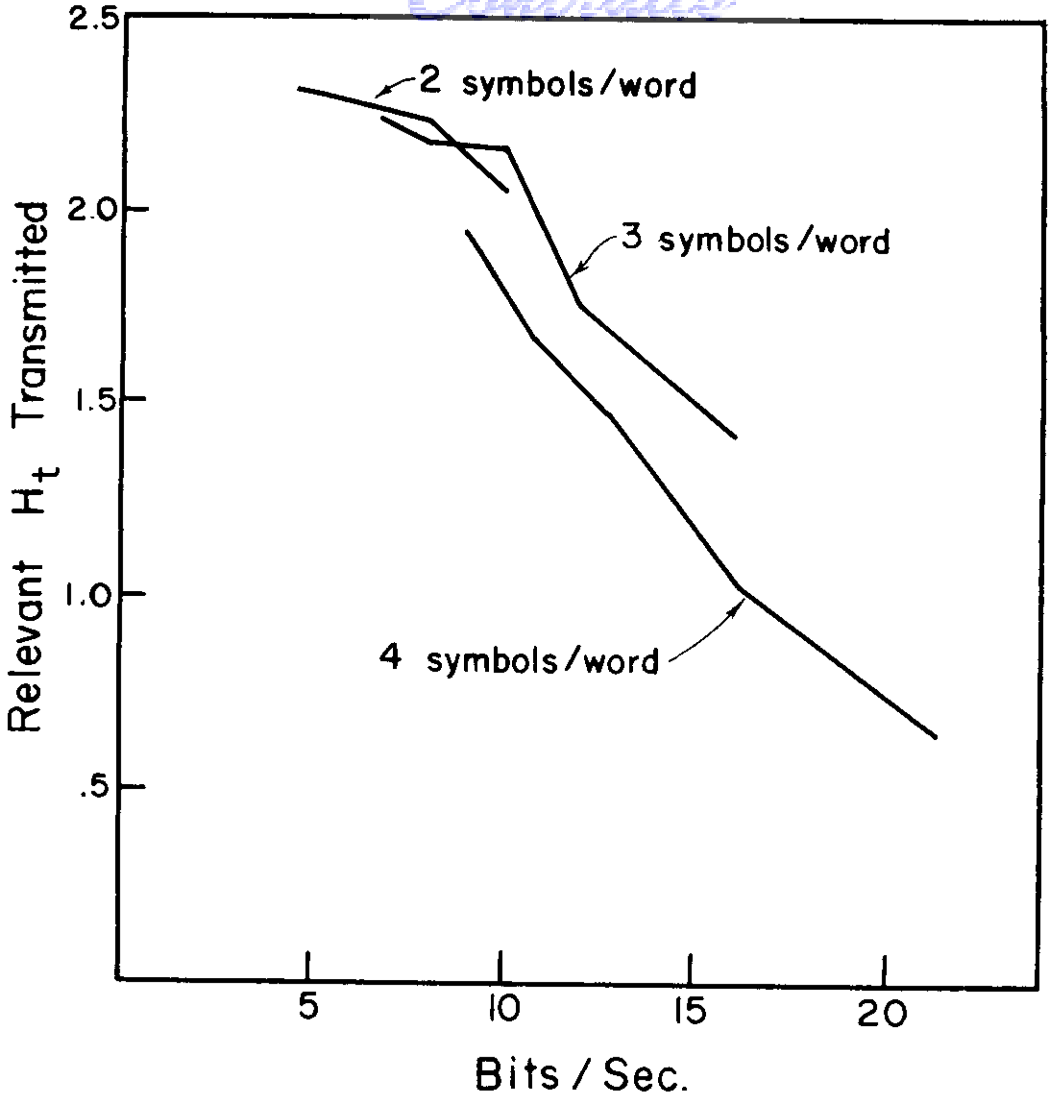

Fig. 5. Relevant information transmitted as a function of the presented stimulus information per second.

given as n(n-1), or the number of possible permutations of the elements to be paired. Postulating the use of "immediate memory" by S, he further predicted that the actual number of comparisons made was n(n-1)/2, since S did not repeat comparisons already made.

The present task also required S to make paired comparisons within each word and sum these for each message. Using the formula, n(n-1)/2, the 2-, 3-, and 4-symbol words should be in the ratio of 1:3:6. However, due to the restrictions imposed upon word construction, this does not give the best fit. In constructing words in this experiment, multiple identities were not used and only a single pair could occur in any one word. Having this knowledge, S's scanning behavior was greatly facilitated. If in scanning a word a pair occurred on the first comparison, no further comparisons were necessary. Since the occurrence of pairs was randomized thoroughly with respect to both frequency and location, the best estimate of the number of paired comparisons made by S is the mean of the possible number of comparisons to be made under each condition. In this study, the means are 1, 2.5, and 4.75, respectively, for the 2-, 3-, and 4-symbol word conditions. With these values, the mean numbers of paired comparisons were calculated for each condition and plotted against relevant information transmission scores. These data are shown in Fig. 6. The plot gives considerable support to the notion that in tasks of this type the average number of pairs to be examined is a meaningful dimension along which to appraise task difficulty. The three load conditions appear almost as a single function with the exception of three points. These points correspond to the highest rate conditions with the 2- and 3-symbol words. There is undoubtedly some limiting rate beyond which the average number of comparisons will not predict the full extent of error. The wide divergence of

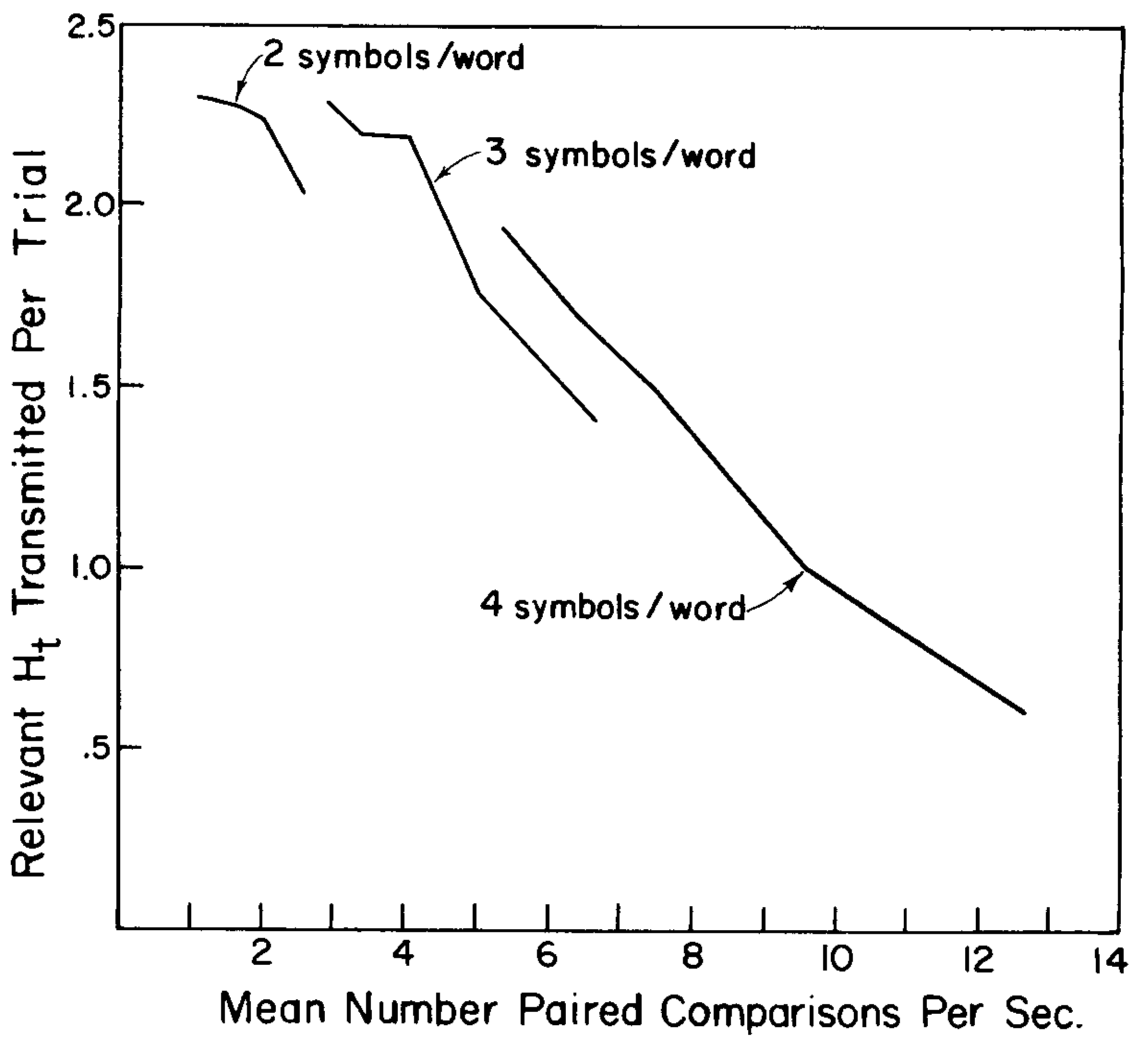

Fig. 6. Relevant information transmitted per trial as a function of mean number of paired comparisons per second.

conditions seen in Fig. 3 and 4 is not present and, in this sense, considering the task from a paired-comparison viewpoint takes full account of the interaction between conditions.

A major difference is apparent if one relates the data in Fig. 5 to the findings of Alluisi, Muller, & Fitts (1957) in their study of information-processing behavior. The earlier study investigated verbal and motor responses to coded stimuli, and it was concluded that at equal rates of information presentation the rate of information transmission was greater for the larger number of alternative stimuli and the slower rate of presentation. In Fig. 5 the area of overlap between the 3- and 4-symbol conditions indicates that when task difficulty was comparable, the amount of relevant information transmitted was lower in the higher load, slower rate conditions. In the present case, it should be noted, S was not required to transmit all the information presented, but selected portions of it. It may be that, in situations of selective discrimination, performance is improved by presenting lower information content messages at higher rates.

SUMMARY

The present investigation dealt with the effects of various levels of speed and load stress on performance in an information-processing task.

The task selected for study required the discrimination and reporting of pairs of identical symbols within symbol matrices of various sizes. Nomenclature from information theory was used to describe the stimulus situation. Each matrix was referred to as a "message." Messages were composed of "words" (rows) and words consisted of symbols (cells). A factorial design was used, including three levels of word length and five levels of rate of presentation. Each of 10 Ss served under all conditions, and performance was assessed in terms of both error scores and a derived information measure.

The results of the study indicated that both speed and load increases seriously impair performance. Nonparametric tests of significance between conditions gave significance at $P < .001$ for all conditions except the shortest word length and the slowest rate. An attempt to fit the data to an objective specification of the stimulus situation in information terms failed to provide a completely satisfactory function. A plot of information scores against the mean number of paired comparisons per second yielded a single function which described all conditions and accounted for the high interaction component in the data.

It was concluded that (a) specification of the stimulus situation alone does not adequately explain performance, and (b) the use of information measures in tasks of this type adds little or nothing to such explanations. A far more pertinent and fruitful consideration in analyzing such tasks is in terms of the actual operations S is required to perform in responding to the situation, in this case, the number of paired comparisons made.

REFERENCES

1. Alluisi, E.A., Muller, P. F., and Fitts, P. M. "An Information Analysis of Verbal and Motor Responses in a Forced-Paced Serial Task." Journal of Experimental Psychology. Vol. 53. 1957. pp. 153-158.

2. Conrad, R. "Speed and Load Stress in a Sensorimotor Skill." British Journal of Industrial Medicine. Vol. 8. 1951. pp. 1-7.

3. Conrad, R. "Some Effects on Performance of Changes in Perceptual Load." Journal of Experimental Psychology. Vol. 49. 1955. pp. 313-322.

4. Mackworth, J. F., and Mackworth, N. H. Temporal Irregularity in a Multi-source Task. Bulletin No. APU 264/57. Medical Research Council, Cambridge, England. 1957.

5. Mackworth, N. H., and Mackworth, J. F. "The Overlapping of Signals for Decisions." American Journal of Psychology. Vol. 69. 1956. pp. 26-47.

6. Sidorsky, R. "An Experimental Study of the Human as a Comparator of Visual Symbols." Unpublished doctor's dissertation, The Ohio State University, Columbus, Ohio. 1954.

7. Siegel, S. Nonparametric Statistics for the Behavioral Sciences. McGraw Hill Book Company, Inc., New York, N. Y., 1956.

8. Wagner, R. C., Fitts, P. M., and Noble, N. E. Preliminary Investigations of Speed and Load as Dimensions of Psychomotor Tasks. AFPTRC Research Report No. 54-45. Air Force Personnel Training and Research Center, Lackland Air Force Base, Texas. 1954.