NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.
Best Available Copy
HUMAN DATA PROCESSING LIMITS IN DECISION MAKING

John R. Hayes

Technical Documentary Report No. ESD-TDR-62-48

July 1962

Operational Applications Laboratory
Deputy for Technology
Electronic Systems Division
Air Force Systems Command
L. G. Hanscom Field, Bedford, Mass.

Project 2806, Task No. 280603
Requests for additional copies by Agencies of the Department of Defense, their contractors, and other government agencies should be directed to the:

Armed Services Technical Information Agency
Arlington Hall Station
Arlington 12, Virginia

Department of Defense contractors must be established for ASTIA services, or have their 'need-to-know' certified by the cognizant military agency of their project or contract.

All other persons and organizations should apply to the:

U. S. DEPARTMENT OF COMMERCE
OFFICE OF TECHNICAL SERVICES,
WASHINGTON 25, D. C.
PSYCHOLOGY

PERSONNEL
AL GUNNERS
BAD CASUALTIES
TRAFFIC CONTROLLERS
BOMBERS
Nuclear specialists
AVIATION PERSONNEL
MISSILE OPERATORS
ARTILlE
LOCAL WEAPONS CASUALTIES
DEAD
MUNITIONS
POLICE PERSONNEL
NAVY CONTROL OPERATORS
NAVY PERSONNEL
ENGINEERING PERSONNEL
AGS
IIC RIFLES
PHYSICISTS
RESEARCH PERSONNEL
ENGINEERING PERSONNEL
2095 GUNNERS
2096 HUMAN ENGINEERING
2097 CYBERNETICS
2098 INDUSTRIAL TRAINING
2099 RADAR TRAINING
2100 TEXTBOOKS
2101 TRAINING
2102 TRAINING DEVICES
2103 TRAINING FILMS

SOCIAL SCIENCES

ANTHROPOLOGICAL DATA
ANTHROPOLOGY
ARCHAEOLOGY
ARCHITECTURE
ARCHITECTURAL DATA
ARCHITECTURAL DESIGN
ARCHITECTURAL EFFICIENCY
ARCHITECTURAL EFFECTIVENESS
ARCHITECTURAL EXHIBITIONS
ARCHITECTURAL HANDBOOKS
ARCHITECTURAL INDEXES
ARCHITECTURAL LABORATORY ANIMALS
ARCHITECTURAL MANUSCRIPTS
ARCHITECTURAL MEASUREMENT
ARCHITECTURAL MEASUREMENTS
ARCHITECTURAL SYMPOSIA
ARCHITECTURAL TEST EQUIPMENT
ARCHITECTURAL TEST METHODS
ARCHITECTURAL THEORY

ASPECT TERMS

ACCEPTABILITY
AIR FORCE RESEARCH
ANALYSIS
BIBLIOGRAPHY
CATALOGS
CLASSIFICATION
COAST GUARD RESEARCH
CONFERENCES
DATA
DESIGN
INTERPRETATION
EFFECTIVENESS
EXPERIMENTAL DATA
FACTOR ANALYSIS
HANDBOOKS
INDEXES
INSTRUCTION MANUALS
LABORATORY ANIMALS
MAN
MEASUREMENT
MILITARY PUBLICATIONS
MILITARY REQUIREMENTS
NAVAL RESEARCH
PRIMATES
RATS
RELIABILITY
REPORTS
SCIENTIFIC REPORTS
SELECTION
STANDARDS
STANDARDS
STATISTICAL DATA
STATISTICS
TEST EQUIPMENT
TEST METHODS
TEST THEORY
HUMAN DATA PROCESSING LIMITS IN DECISION MAKING

John R. Hayes

Technical Documentary Report No. ESD-TDR-62-48

July 1962

Operational Applications Laboratory
Deputy for Technology
Electronic Systems Division
Air Force Systems Command
L. G. Hanscom Field, Bedford, Mass.

Project 2806, Task No. 280603
HUMAN DATA PROCESSING LIMITS IN DECISION MAKING

Abstract

Four experiments are described in which subjects were required to choose among alternatives on the basis of two, four, six, or eight relevant facts. Both decision quality and decision time were measured. Presenting more than four facts caused a decrease in decision making efficiency.

Reviewed and approved for publication

HERBERT RUBENSTEIN, Chief
Information Presentation Division
Operational Applications Laboratory

WILLIAM V. HAGIN, Colonel, USAF
Director
Operational Applications Laboratory
HUMAN DATA PROCESSING LIMITS IN DECISION MAKING

J. R. HAYES

INTRODUCTION

Currently, the name "decision making" is applied to a very large group of behaviors which range in complexity from predicting which of two lights is about to light (1, 2) to establishing plans for conducting a war. This report will be concerned with decisions similar to the decisions involved in choosing which one of a number of cars to buy or which one of several apartments to rent. Most usually in such decisions, the alternatives will differ from one another in several characteristics, and these differences must be taken into account simultaneously in making the choice. For example, in choosing among alternative apartments, one may consider cost, size, appearance, convenience of location, quality of neighborhood, and possibly a number of other characteristics. The difficulty in making such decisions arises in trading the advantages of an alternative in some characteristics against its disadvantages in other characteristics. Such decisions might be described as multi-dimensional judgments.

It is commonly assumed that the more relevant data one takes into account in making a decision, the better that decision will be. It
is clear, however, that as one takes more relevant characteristics into account in comparing alternatives, the opportunities for confusion increase. If confusion were to increase rapidly enough as the number of characteristics increased, it is conceivable that decision makers would perform better if some of the relevant data were eliminated. The experiments reported below concern the manner in which decision making performance varies with the number of characteristics presented in a complex decision problem. Two aspects of performance were measured -- decision "quality" (to be defined explicitly below) and decision time.

The Decision Task

In all of the experiments, the subjects, naval enlisted men made decisions in a simulated military problem. The task was to dispatch one of several airplanes to investigate a reported submarine sighting. The decisions were to be based on information contained in a data matrix of the form shown in Fig. 1. The top row lists the alternative airplanes and the left hand column lists several characteristics of the alternatives. The number of alternative airplanes varied from two to eight depending on the experiment. In all experiments, problems involving two, four, six, and eight characteristics were presented.

In all of the experiments, the characteristics used were chosen from the same list of eight characteristics, in the order listed in Fig. 1 these were: the quality of the pilot, the airplane's armament (measured as the percentage of full ammunition load on board), the
capability of the airplane to make visual contact with a surface object, the speed of the airplane, the delay until the airplane could take off, the quality of the airplane's radar, the distance of the airplane's base from the target, and the time available to search for the target once the plane had reached the target area. The characteristics listed above are not necessarily the ones which military experts would consider most important, but are sufficiently relevant to provide the experimental task with adequate face validity for the subjects employed.

For a particular airplane, each characteristic could take on any one of the eight values listed in Table 1. For example, the speed of an airplane could take on values between 375 M.P.H. and 165 M.P.H. in 30 M.P.H. steps and the quality of the pilot could take on values from E (excellent) to VP (very poor). The values of each characteristic could be ranked unequivocally from one for the "best" or most desirable value to eight for the "worst" or least desirable value. These rankings are also shown in Table 1.

The eight characteristics do not share a common unit of measurement, nor does an increasing numerical value necessarily indicate an increase in desirability. For example, higher numbers are more desirable for "speed" but less desirable for "delay". These properties of the set of characteristics make it unlikely that the subject will adopt a simple arithmetic rule for arriving at decisions and, hence, will help to insure that the decisions will involve genuinely multidimensional judgments.
It is also important in obtaining multi-dimensional judgments that the range of relative weights of the characteristics not be too great. An example will help to make clear what is meant by "relative weights". Suppose that for a given subject an airplane with the best speed (375 M. P. H.) and the third best delay (4 min.) was judged equal to an airplane with the second best speed (345 M. P. H.) and the best delay (0 min.). For this subject, a difference of one rank unit of speed balanced a difference of two rank units of delay. Delay, then, had a relative weight of 50% with respect to speed. If delay had had a relative weight as low as 10% with respect to speed, then the whole range for delay, from one rank unit to eight rank units, would not be sufficient to balance a single rank unit difference in speed. In a decision based on these two characteristics, the subject would have to take delay into account only when the alternatives had identical speeds.

The ranges of values for the various characteristics shown in Table I were chosen, on the basis of experience with preliminary tests, to yield a small range of relative weights. For these and other reasons to be discussed below, it is felt that the range of relative weights has been kept sufficiently small for the purposes of the present study.

Measures of Performance

Data were collected on decision time and on two measures of decision quality: the information transmitted from the data matrix to the decisions and the mean grade of the decisions. The two measures
of quality were intended to supplement one another.

The decision time was the interval in seconds between the moment at which the subject was told to expose the data matrix and the moment at which he announced his decision.

The information transmitted from the data matrix to the decisions, T(M,D), was calculated by the technique of McGill (6) from the number of subjects choosing each alternative. (See Appendix B for a discussion of this technique.) T(M, D) was used as a measure of the quality of decisions because its magnitude reflected the degree to which the decisions were determined by the data presented.

While large values of T(M, D) would indicate that the subjects' choices were strongly influenced by the data presented, a question could be raised as to whether the choices were "good" ones or not. It was decided, therefore, to supplement the information transmission measure

Footnote 1

Ideally, one should measure the information transmitted from the data matrix to the decisions for each subject rather than for the group of subjects. As Garner (3) has shown, individual differences may become more marked as the complexity of the judgmental situation increases. The presence of such an effect in the present data would cause a progressive underestimation of the predictability of the decisions as the number of characteristics was increased. Unfortunately, practical limitations made it impossible to measure transmitted information separately for each subject.
with a measure of the "goodness" of the choices. The measure used here is called "the mean grade" of the decisions.

The mean grade was computed as follows: A grade from zero to 100 was assigned to each alternative of each problem by a procedure to be described below. The mean grade for a set of decisions was the average grade of the alternatives chosen in those decisions. The following procedure was used to assign grades to alternatives:

Consider the data matrix shown in Fig. 1. Each value of each characteristic in the matrix was assigned a rank from one for "best" through eight for "worst" as is shown in Table 1. Figure 2 shows the ranks so assigned together with the sums of ranks for each alternative airplane. The alternative with the smallest sum of ranks in each problem was designated "best" and given a grade of 100. The alternative with the largest sum of ranks was designated "worst" and given a grade of zero. Alternatives between these two were given grades proportional to the difference in sum of ranks between the given alternative and the "worst" alternative. The measure of the quality of a set of decisions was the mean grade of that set of decisions.

It is clear that the sum of ranks criterion ignores the interactions among the characteristics which the subjects might try to take into account in making their decisions.

Yntema and Torgerson (8) have shown however, that the effect of ignoring such interactions in decisions of the type described here are not likely to be serious.
The sum of ranks criterion "works" in the sense that, on the average, it orders the alternatives in the same way that the subjects do. In the four alternative problems of Experiment I, for example, the subjects chose the "best" alternative 52.1% of the time, the "second best," 29.4% of the time, the "third best," 12.2% of the time, and the "worst," 6.3% of the time. The same tendency is reflected in the behavior of individual subjects. Every subject in all of the experiments reported below chose the "best" alternative more frequently than any other.

Experiment I

Subjects: 16 naval enlisted men with GCT scores of 50 or more.

Materials: The materials consisted of 240 data matrices of the general form shown in Fig. 1 each printed on an 8" by 10 1/2" sheet of white paper. The 240 matrices were made up of 12 families of ten four-alternative matrices and 12 families of ten eight-alternative matrices.

The construction of a family of matrices involved the following steps:

1. The list of eight characteristics was divided into four pairs, e.g. speed-pilot, distance-delay, etc.

Footnote

1Details of the methods used to construct the materials for Experiments 1, 2, and 3 are given in Appendix A.
2. Alternatives were chosen to be best, second best, third best, and fourth best in terms of the sum of ranks criterion, (e.g. airplane 5 might be best in all ten problems, airplane 3 second best, etc.). In the four alternative problems, there was only one fourth best alternative, but in the eight alternative problems, there were five fourth best alternatives all with the same sum of ranks.

3. Values were assigned to the characteristics in such a way that the sum of ranks for each alternative over each pair of characteristics was one less than for the next best alternative, e.g. the sum of ranks over the speed-pilot pair might be 8 for the best alternative, 9 for the second best, 10 for the third best, and 11 for the one or five fourth best alternatives.

4. The seven standard matrices in a family were constructed by entering one or more of the pairs of characteristics into a matrix. The four two characteristic matrices each consisted of one such pair. The four, the six, and the eight, characteristic matrices consisted respectively of two, three, and four, such pairs. The four, six, and eight characteristic problems, then, may be viewed as a combination of various of the two characteristic problems.

5. The three reduced information matrices, consisting of one four, one six, and one eight characteristic matrix, were the same as the corresponding standard problems with one modification. One of the pairs of characteristics in the standard matrix was modified so that the sum of ranks over that pair was the same for all alterna-
tives. Thus, in the reduced information problems, one of the pairs was not useful in locating the best alternative.

To summarize, in the standard problems, the advantage of each alternative over the next best alternative was proportional to the number of characteristics involved. In the two, four, six and eight, characteristic problems, the advantage was, respectively, one, two, three, and four rank units. In the reduced information problems, the advantage was one rank unit less than in the corresponding standard problems. Table 2 shows the number of standard problems and the number of reduced information problems with each number of alternatives and each number of characteristics.

Procedure

Before testing, each subject was instructed as follows:

"This is an experiment in decision making. You are based at a shore station and you receive reports of radar sightings of submarines. It is your job to dispatch a single plane to search the area where the sighting occurred. To do this you have to decide which one of the available planes is best for the assignment."

"In each problem, you will have either four or eight planes from which you must choose. In making your decision you will have to consider several factors which describe the planes."

At this point the subject was shown a data matrix listing the eight characteristics. Each was explained to him in detail. The
subject was then asked to paraphrase the explanations. If the subject's paraphrasing was judged unsatisfactory, the explanation was repeated.

The instructions then continued as follows:

"The problems will be presented to you on these sheets." (E shows S a data matrix), " and will be placed face down in front of you. When I give the signal turn the problem over and start work. When you have finished, tell me which plane you have chosen

"Each day's test will consist of 20 problems. In each problem, you should try to make the best decision possible; accuracy is the most important thing."

After instruction, the subjects solved a set of 20 practice problems which included all of the problem types shown in Table 2. Each problem type was represented with frequencies proportional to those shown in Table 2. The subjects then solved 20 problems a day in each of 12 daily experimental sessions. No session included more than one problem from any problem family.

In no case were the subjects told whether or not their decisions were "correct," nor were they told how to weigh the various characteristics. All subjects were tested in individual sessions.

RESULTS

Figure 3 shows that decision time increased markedly as the number of characteristics was increased both for four and for eight alternative problems. The significance of differences in decision
time between problem types was tested by two-tailed sign test (7) with a 5% significance level. All differences among the four alternative standard problem types and among the eight alternative standard problem types were significant except for the difference between the four and six characteristic problems with eight alternatives. Decision time was significantly greater for all eight alternative standard and reduced information problem types than for the corresponding four alternative problem types. The reduced information problem types did not differ significantly from the standard problem types except for the eight alternative four characteristic problems.

While decision time increased as the number of characteristics was increased, decision quality did not. Figure 4 shows that the information transmitted from the data matrix to the decisions did not increase as the number of characteristics was increased. In fact, a fairly marked drop may be observed between two and four characteristics for the eight alternative problems.

Similarly, Figure 5 shows that the mean grade of the decisions did not increase as the number of characteristics was increased. The significance of differences in mean grade was tested by two-tailed sign tests (7) with a 5% significance level. Mean grades for the four alternative standard and reduced information problem types were all significantly greater than for the corresponding eight alternative problem types except for the standard six characteristic problems. None of the differences among the standard four alternative problem
types was significant. The only significant difference among the standard eight alternative problem types was the decrease in mean grade from two to four characteristics.

DISCUSSION

The failure to find any increase in decision quality as the number of characteristics increased was unexpected. It lead the experimenters to test two hypotheses which might account for the result. The first was that the measures of decision quality were insensitive to real changes in the quality of the decisions. This hypothesis was tested by comparing the reduced information problems with the standard problems. It was reasoned that the quality of decisions should be higher for the standard problem than for the corresponding reduced information problems since the standard problems contained a higher proportion of useful data than did the reduced information problems. As Figures 4 and 5 show, the reduced information decisions were in all cases lower in quality on either measure than the corresponding standard decisions. The difference in mean grade between the reduced information decisions and the standard decisions was significant at the 5% level by two-tailed sign tests (7). The hypothesis that the results of the experiment could be attributed to insensitivity of the measures of quality was therefore rejected.

The second hypothesis was that more data could not lead to better decisions in the particular problems chosen for study. One
might argue, for example, that since the information in each pair of characteristics was in a sense redundant with the information in every other pair, four pairs of characteristics could give the subject no more information than one pair. Therefore, one would expect no improvement in decision performance as the number of characteristics was increased.

The test of this hypothesis involved demonstrating that more data could, in fact, be used to produce better decisions. As noted above, each pair of characteristics in each standard eight characteristic problem had been presented separately as a two characteristic problem. The four two characteristic decisions were used to construct a synthetic eight characteristic decision by the use of a simple voting rule. That is, the alternative chosen in the synthetic decision was the alternative most often chosen in the four two characteristic decisions. When two or more alternatives were tied, one of them was chosen by a random process.

Table 3 shows that the mean grades over all 16 subjects for the synthetic decisions were higher than for the standard two and the standard eight characteristic decisions. The columns labeled n in Table 3 show the numbers of subjects for whom the mean grade for the synthetic decisions was greater than the mean grade for the standard decisions. The differences between synthetic and standard decisions were all significant beyond the 5% level by two-tailed sign tests (7). Since extra data, properly used, could lead to better de-
cisions, the second hypothesis was also rejected.

A simple two factor theory is proposed to account for the results with respect to decision quality\(^1\). It is assumed that the ability of subjects to identify good alternatives depends on two factors. First, it depends on the differences in value among the alternatives. (Value is crudely represented in this experiment by sum of ranks.) Other things being equal, the greater the difference in value, the more likely it is that the subjects will choose the best alternative. Second, it depends on the subject's sensitivity, that is on his ability to judge which of several alternatives has the greatest value. Other things being equal, the subject's sensitivity decreases as the number of characteristics he takes into account increases.

Several of the experimental results are consonant with the theory. The differences between standard problems and reduced information problems of the same size can be accounted for by the first factor. As Figures 4 and 5 show, the reduced information problems, characterized by small sums of ranks differences, all yield decisions of lower quality than do the corresponding standard problems.

---

Footnote

\(^1\)This theory will be developed more fully in a later publication. This simple account is given here in an attempt to make the presentation more orderly.
One could account for the lack of differences among standard problems by assuming that as the number of characteristics was increased, the advantage of the linearly increasing differences in value was just offset by a linearly decreasing sensitivity of the subjects. The assumption that sensitivity decreases linearly as the number of characteristics is increased is ad hoc at this point. However, as will be shown in a later publication, this assumption is useful in describing other aspects of the data, e.g. the superiority of the synthetic decisions, and the tendency for the quality of the reduced information decisions to increase as the number of characteristics is increased.

Experiment 2

Apparently the subjects of Experiment 1 were not able to use extra data efficiently enough for it to yield an improvement in decision quality. It seemed reasonable that this inefficiency might be due to lack of training and that training of the subjects in decision making might reduce it. Experiment 2 was performed to see whether or not the decisions of subjects trained in decision making would improve in quality as the number of characteristics was increased.

Subjects: The subjects were 16 naval enlisted men with G. C. T. scores of 50 or more.

Materials: The materials were the same as those used in Experiment 1.

Procedure

The subjects were arbitrarily divided into two groups of eight.
subjects each, Group A and Group B, and the problems were divided into two sets, Set 1 and Set 2. Set 1 contained six of the 12 four alternative problem families and six of the 12 eight alternative problem families. Set 2 contained the remaining problem families. Thus, each set contained 12 complete families of ten problems each or 120 problems. Group A was trained on Set 1 and tested on Set 2. Group B was trained on Set 2 and tested on Set 1.

The subjects solved 20 problems a day in 13 daily sessions. Session 1 was devoted to the 20 practice problems used in Experiment 1. Sessions 2 through 7 were devoted to the training problems and sessions 8 through 13, to the test problems. The initial instructions and the procedures during test sessions were the same as those used in Experiment 1. During sessions 1 through 7, however, the subject was told after each decision whether his choice was correct or incorrect. The experimenter accepted as correct only the alternative designated "best" by the sum of ranks criterion. Whenever the subject made an incorrect choice, he was required to choose again until he made the correct choice.

RESULTS

Only data from the test sessions are reported here. The significance of differences among problem types was in all cases tested by two-tailed sign tests (7) with a 5% significance level.

The results of Experiment 2 closely parallel those of Experiment 1. Figure 6 shows that decision time increased markedly...
as the number of characteristics was increased both for the four and for the eight alternative problems. All differences among four and among eight alternative standard problem types were significant with two exceptions. These were the differences between six and eight characteristic problems with four alternatives and the difference between four and six characteristic problems with eight alternatives. Decision time was significantly greater for each eight alternative problem type than for the corresponding four alternative problem type. The reduced information problem types did not differ significantly from the corresponding standard problem types except for the eight alternative six characteristic problems.

The results concerning decision quality were also very similar to those found in Experiment 1. Figure 7 shows that the information transmitted from the data matrix to the decisions did not tend to increase as the number of characteristics was increased. Figure 8 shows that while the mean grades were all somewhat higher in Experiment 2 than in Experiment 1, they did not tend to increase as the number of characteristics was increased. None of the differences in mean grade among the standard four or among the standard eight alternative problem types was significant. Mean grades for the four alternative problem types were significantly greater than for the eight alternative problem types. The reduced information problems yielded significantly lower grades than the standard problems.

Synthetic decisions were constructed in the manner discussed
under Experiment 1, from the two characteristic decisions made during the test sessions. As Table 4 shows, the mean grades for the synthetic decisions were higher than for the two or eight characteristic standard decision. Only the differences between synthetic and two characteristic decisions were significant, however. The columns labeled \( n \) in Table 4 show the numbers of subjects, out of 16 for whom the mean grade for the synthetic decisions was greater than for the standard decisions.

**DISCUSSION**

With the exception that the mean grades for decisions in Experiment 2 were somewhat higher than for those in Experiment 1, the results of the two experiments were almost exactly parallel. Training in decision making, to the extent that it was provided in Experiment 2, did not appear to change the relations found in Experiment 1.

**Experiment 3**

Experiment 3 was performed to investigate the effects of limiting decision time on decision quality.

**Subjects:** The subjects were eight naval enlisted men with G.C.T. scores of 50 or more.

**Materials:** The materials were the same as those used in Experiment 1 except that only one of the four two characteristic problems in each family was used. The one problem to be used was selected by a random process. Thus, the materials consisted of 24
families of seven problems each or a total of 168 problems.

**Procedure**

Rather than being given unlimited time to solve the problems as in Experiment 1, the subjects were told, "There will be a ten second time limit on each problem. When the ten seconds are up, you must give me your decision immediately without further study of the problem."

The subjects solved 14 problems a day in 13 daily sessions. Session 1 was devoted to practice problems and Session 2 through 13 to test problems.

In all other respects the instructions and procedures were the same as in Experiment 1.

**RESULTS**

Figure 9 shows that the information transmitted from the data matrix to the decisions tended to decrease in the standard problems as the number of characteristics was increased. Figure 10 shows that the mean grade of the standard decisions tended to decrease as the number of characteristics was increased in the eight alternative case but not in the four alternative case. The decreasing trend was found significant at the 5% level by non-parametric trend test (5) in the eight alternative case, but not significant in the four alternative case. The differences in mean grade between the four and eight alternative standard problems were significant at the 5% level by two-tailed sign tests (7) for the four, six, and eight characteristic problems.
but not for the two characteristic problems.

**DISCUSSION**

With unlimited decision time as in Experiment 1 and 2, increasing the number of characteristics did not increase the decision quality. With limited decision time as in Experiment 3, increasing the number of characteristics could actually decrease decision quality.

**Experiment 4**

The problem material used in the first three experiments was especially constructed to facilitate analysis. In particular, two constraints were placed on the standard problems. First, the difference between each alternative and the next best alternative was fixed at one rank unit per pair of characteristics. Second, the problems had uniform structure; that is, they were constructed by adding together from one through four two characteristic problems all of which had the same best alternative, the same second best alternative, etc. Because of these constraints one may question the generality of the obtained results. Therefore, they were removed in Experiment 4.

**Subjects:** The subjects were 14 naval enlisted men with G. C. T. scores of 50 or more.

**Materials:** The materials consisted of 300 two alternative problems equally divided among the two, four, six, and eight, characteristic types. That is, there were 75 problems of each type. The 75 eight characteristic problems were selected from a large set.
of problems which had been constructed by assigning values at random to each of the eight characteristics. The selection was made randomly with the restriction that the rank unit difference between the alternatives be either one, two, three, five, or eight, and that each of these differences be represented in 15 problems.

Two characteristics were subtracted at random from each eight characteristic problems to generate the 75 six characteristic problems. In the same way, the six characteristic problems were used to generate the four characteristic problems and these in turn were used to generate the two characteristic problems. No effort was made to control the rank unit differences between the alternatives in the two, four, and six characteristic problems. The columns labeled h in Table 5 show the numbers of problems with each rank unit difference for each number of characteristics.

Procedure

The instructions were the same as in Experiment 1 with the exception that the subjects were told that there would be two alternative airplanes in each problem rather than four or eight. The subjects solved 20 problems a day in each of 16 daily sessions. Session 1 was devoted to practice problems and sessions 2 through 16 to test problems.

RESULTS

Figure 11 shown decision time as a function of the rank unit difference between alternatives. Number of characteristics is the
Mean decision times were computed only for points representing three or more problems. As in Experiments 1 and 2, decision time increased markedly as the number of characteristics was increased. Decision time appeared to decrease somewhat as the rank unit difference between the alternatives was increased, especially in the four and six characteristic problems. The effect was much less marked, however, than for changes in the number of characteristics. This result confirms the finding of Experiments 1 and 2 that decision time was about the same for standard and reduced information problems with the same number of characteristics.

The columns labeled P in Table 5 give the observed probabilities of choosing the "better" of the two alternatives. Probabilities were computed only for groups of three or more problems. In Fig. 12, these probabilities were plotted on a cumulative normal scale (4) against the rank unit difference per characteristic. It is apparent that the points lie approximately on a single straight line. Thus, when the rank unit difference per characteristic was held constant the probability of choosing the "better" alternative did not change as the number of characteristics was increased. Since the mean grade in a two alternative decision is just the probability of choosing the "better" alternative multiplied by 100, it is clear that this result confirms and extends the results for the standard problems found in Experiments 1 and 2. Previously, the result had been obtained for problems with
uniform structure in which the rank unit difference per characteristic had been fixed at 0.5. In the present study, the result was obtained for problems with random structure and a wide range of rank unit differences per characteristic.
APPENDIX A

Method used in constructing the 240 data matrices.

Step 1. Construction of 12 standard 8-alternative-8-characteristic problems and 12 standard 4-alternative-8-characteristic problems.

a. Best, second best, and third best alternatives were chosen for each problem by a random process. The remaining 5 alternatives in the eight alternative problems and the remaining one alternative in the four alternative problems were designated "worst."

b. The eight characteristics were divided into four pairs of two characteristics each. For six of the four-alternative problems and six of the eight-alternative problems the pairings were speed-pilot, distance-delay, armament-search, and contact-radar. For the remaining problems, the pairings were speed-distance, pilot-delay, contact-search, and armament-radar.

c. Within each problem a quality level was chosen for each pair of characteristics. This was done by choosing a random number between 7 and 13 for each of the four pairs of characteristics within each problem. The number chosen, S, represented the sum of ranks over the two characteristics of the pair to be assigned for the worst alternatives. The best, second best, and third best alternatives were assigned sums of ranks of S - 3, S - 2, and S - 1 respectively.
The quality level was varied in this way from problem to problem and within each problem so that the subjects could not learn to identify the best alternative by an absolute judgment of quality. With the procedure described above, the worst alternative in one problem could be of better quality than the best alternative in the next problem.

Table 6 illustrates the procedures up to this point for the data matrix shown in Fig. 1.

d. Each characteristic was then assigned a rank for each alternative. The ranks were assigned at random with the restriction that for each pair of characteristics the ranks add to the sum designated in c above. The sums of ranks used in constructing the data matrix in Fig. 1 are given in Table 6. The first cell of Fig. 6 shows that the sum of ranks assigned for speed and pilot was 12 for the first alternative. Therefore a random choice was made among the pairs of ranks 4-8, 5-7, 6-6, 7-5, and 8-4. Table 7 shows that the choice was 6 and 6.

e. The order in which the characteristics were listed in the matrix was then randomized. The procedure to this point is illustrated in Fig. 2.

f. The assigned ranks were then translated into category values by the use of Table 1. Figure 1 illustrates a completed data matrix.
Step 2. The construction of two alternative, four alternative, and six alternative, standard problems.

a. In d of Step 1 above, a matrix such as that shown in Table 8 was generated for each eight characteristic problem. The procedures of Step 2 start with these matrices. Each eight characteristic matrix was used to generate four two characteristic matrices simply by placing each pair of characteristics in a separate matrix. Four and six characteristic matrices were generated by selecting respectively two and three of the four pairs of characteristics at random. Thus, each eight characteristic matrix was used to generate six new matrices.

b. The order in which the characteristics were listed was randomized and the assigned ranks were translated into category values by the use of Table 1.

c. In placing the data in the matrices, no gaps were left. For example, in two characteristic problems, the two characteristics were placed in the top two rows of the matrix, the remainder of the matrix remaining blank.

Step 3. The construction of four, six, and eight characteristic reduced information problems.

a. Each four, six, and eight characteristic standard problem was used to generate a corresponding reduced information problem. One pair of characteristics in each standard
problem was selected at random for modification.

b. The modification consisted in making the best, second best, and third best, alternatives within the selected pair equivalent to the worst alternatives. This was done by making the sums of ranks selected in c of Step 1 for the three favored alternatives all equal to the sum of ranks for the worst alternatives. Thus, if the distance-delay pair of Table 6 had been selected for modification, alternatives 2, 5, and 8 would then be assigned a sum of ranks of 11.

c. New values of the selected pair of characteristics were assigned to the modified alternatives by applying procedures d and f of Step 1.

d. The order in which the characteristics were listed in the modified matrix was then randomized again, so that the order of listing in the reduced information problem was different from that in the corresponding standard problem.
The information transmitted from the data matrix to the
decisions, $T(M,D)$, was measured as follows: First, the number
of subjects choosing each alternative was tabulated for each problem
as shown in Table 8. Since each subject made only one decision in
each problem, $F_1 = F_2 = \ldots = F_n = s$, where $s$ is the number of
subjects.

Second, $T(M,D)$ was computed from the formula $T(M,D) =$

$$H(M) + H(D) - H(M,D)$$

where

$$H(M) = \frac{1}{s} \sum_{i=1}^{n} F_i \log_2 \frac{F_i}{N}$$

and

$$H(D) = \frac{1}{k} \sum_{i=1}^{n} F_i \log_2 \frac{F_i}{N}$$

and

$$H(M,D) = \frac{1}{s} \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{F_{ij}}{N} \log_2 \frac{F_{ij}}{N}$$

For a fuller discussion of this technique see McGill (6).

$T(M,D)$ is used as measure of decision quality because it
reflects the degree to which the decisions are determined by the data
matrix. If the choices were made without reference to the data matrix,
then $T(M,D) = 0$. Increasing influence of the data matrix on the
decisions will yield increasing values of $T(M,D)$ up to a maximum
value of $\log_2 k$, where $k$ is the number of alternatives. Good decision
processes, of course, should be influenced strongly by the data, and,
hence, should yield large values of $T(M,D)$. 
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot</td>
<td>PF</td>
<td>F</td>
<td>FG</td>
<td>P</td>
<td>G</td>
<td>F</td>
<td>FG</td>
<td>P</td>
</tr>
<tr>
<td>Armament</td>
<td>40% 90% 100% 80% 90% 90% 60% 60%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact</td>
<td>VP</td>
<td>FG</td>
<td>P</td>
<td>F</td>
<td>P</td>
<td>VP</td>
<td>PF</td>
<td>FG</td>
</tr>
<tr>
<td>Speed</td>
<td>225</td>
<td>255</td>
<td>165</td>
<td>255</td>
<td>225</td>
<td>195</td>
<td>165</td>
<td>285</td>
</tr>
<tr>
<td>Delay</td>
<td>14</td>
<td>10</td>
<td>12</td>
<td>8</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Radar</td>
<td>E</td>
<td>FG</td>
<td>P</td>
<td>F</td>
<td>GE</td>
<td>PF</td>
<td>G</td>
<td>GE</td>
</tr>
<tr>
<td>Distance</td>
<td>140</td>
<td>140</td>
<td>160</td>
<td>200</td>
<td>180</td>
<td>240</td>
<td>220</td>
<td>240</td>
</tr>
<tr>
<td>Search</td>
<td>31</td>
<td>19</td>
<td>25</td>
<td>13</td>
<td>43</td>
<td>31</td>
<td>19</td>
<td>13</td>
</tr>
</tbody>
</table>

Figure 1
A Typical Data Matrix
Figure 2
Ranks and Sums of Ranks for the Matrix
Shown in Figure 1

<table>
<thead>
<tr>
<th>ALTERNATIVES</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Armament</td>
<td>7</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Contact</td>
<td>8</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>8</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Speed</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Delay</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Radar</td>
<td>1</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>2</td>
<td>6</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Distance</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Search</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>8</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Sum of Ranks</td>
<td>44</td>
<td>36</td>
<td>44</td>
<td>44</td>
<td>32</td>
<td>44</td>
<td>44</td>
<td>40</td>
</tr>
</tbody>
</table>
Figure 3
Mean Decision Times in Experiment 1

40
30
20
10
0
2
4
6
8
CHARACTERISTICS

MEAN DECISION TIME IN SECONDS

8 ALTERNATIVES
REDUCED INFORMATION

8 ALTERNATIVES
STANDARD

4 ALTERNATIVES
REDUCED INFORMATION

4 ALTERNATIVES
STANDARD
Figure 4
Information Transmissions In Experiment 1

TRANSMITTED INFORMATION IN BITS

2 4 6 8
CHARACTERISTICS

8 ALTERNATIVES
STANDARD

8 ALTERNATIVES
REDUCED INFORMATION

4 ALTERNATIVES
STANDARD

4 ALTERNATIVES
REDUCED INFORMATION
Figure 5
Mean Grade of Decisions in Experiment 1

- 4 Alternatives
  - Reduced Information
  - Standard

- 8 Alternatives
  - Standard
  - Reduced Information
Figure 6
Mean Decision Times in Experiment 2

- 8 Alternatives
  - Reduced Information
- 4 Alternatives
  - Standard
  - Reduced Information
Figure 7
Information Transmissions In Experiment 2

- 8 ALTERNATIVES STANDARD
- 8 ALTERNATIVES REDUCED INFORMATION
- 4 ALTERNATIVES STANDARD
- 4 ALTERNATIVES REDUCED INFORMATION

TRANSMITTED INFORMATION IN BITS

CHARACTERICS

35
Figure 8
Mean Grade of Decisions in Experiment 2

- 4 Alternatives
  - Reduced Information
  - Standard

- 8 Alternatives
  - Standard
  - Reduced Information

Mean Grade vs. Characteristics

36
Figure 9
Information Transmissions in Experiment 3

TRANSMITTED INFORMATION IN BITS

CHARACTERISTICS

8 ALTERNATIVES
STANDARD

8 ALTERNATIVES
REDUCED INFORMATION

4 ALTERNATIVES
STANDARD

4 ALTERNATIVES
REDUCED INFORMATION
Figure 10
Mean Grade of Decisions In Experiment 3

- 4 ALTERNATIVES
  - STANDARD
  - REDUCED INFORMATION

- 8 ALTERNATIVES
  - STANDARD
  - REDUCED INFORMATION

MEAN GRADE

CHARACTERISTICS
Figure 11

Mean Decision Times in Experiment 4
Figure 12
Probability of Choosing the "Best" Alternative

PERCENT "BEST" CHOICE

RANK UNITS PER CHARACTERISTIC

2 CHARACTERISTICS - O
4 "   " - ▲
6 "   " - +
8 "   " - △
<p>| Characteristic | “Best” Value | | | | | | “Worst” Value | Unit of Measurement |
|---------------|--------------|---|---|---|---|---|---|---|------------------|
|               | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |                  |
| Speed         | 375 | 345 | 315 | 285 | 255 | 225 | 195 | 165 | Miles per hour |
| Pilot         | E  | GE | G  | FG | F  | PF | P  | VP | Adjective       |
| Delay         | 0  | 2  | 4  | 6  | 8  | 10 | 12 | 14 | Minutes         |
| Radar         | E  | GE | G  | FG | F  | PF | P  | VP | Adjective       |
| Armament      | 100% | 90% | 80% | 70% | 60% | 50% | 40% | 30% | Percent        |
| Distance      | 100 | 120 | 140 | 160 | 180 | 200 | 220 | 240 | Miles           |
| Search        | 55  | 49  | 43  | 37  | 31  | 25  | 19  | 13  | Minutes         |
| Contact       | E  | GE | G  | FG | F  | PF | P  | VP | Adjectives      |</p>
<table>
<thead>
<tr>
<th>Type</th>
<th>4 Alternatives</th>
<th>8 Alternatives</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Standard</td>
<td>Reduced Information</td>
</tr>
<tr>
<td>Number</td>
<td>2 48</td>
<td>--</td>
</tr>
<tr>
<td>of Characteristics</td>
<td>4 12</td>
<td>12</td>
</tr>
<tr>
<td>Number of Characteristics</td>
<td>6 12</td>
<td>12</td>
</tr>
<tr>
<td>Number of Characteristics</td>
<td>8 12</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>84 36</td>
<td>84</td>
</tr>
</tbody>
</table>
TABLE 3

<table>
<thead>
<tr>
<th>Standard Two Characteristics</th>
<th>Standard Eight Characteristics</th>
<th>Synthetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Grade n</td>
<td>Mean Grade n</td>
<td>Mean Grade</td>
</tr>
<tr>
<td>4 Alternatives 76.6 15</td>
<td>77.1 13</td>
<td>82.5</td>
</tr>
<tr>
<td>8 Alternatives 70.4 16</td>
<td>65.8 14</td>
<td>80.3</td>
</tr>
</tbody>
</table>
### TABLE 4

<table>
<thead>
<tr>
<th></th>
<th>Mean Grade</th>
<th>n</th>
<th>Mean Grade</th>
<th>n</th>
<th>Mean Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Two Characteristics</td>
<td>82.2</td>
<td>16</td>
<td>83.6</td>
<td>10</td>
<td>89.5</td>
</tr>
<tr>
<td>Standard Eight Characteristics</td>
<td>76.4</td>
<td>15</td>
<td>79.2</td>
<td>10</td>
<td>86.5</td>
</tr>
</tbody>
</table>
Table 5
Probabilities of Choosing the "Better" Alternative

<table>
<thead>
<tr>
<th>DIFFERENCE IN RANK UNITS</th>
<th>NUMBER OF CHARACTERISTICS</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>P</td>
<td>n</td>
<td>P</td>
<td>n</td>
</tr>
<tr>
<td>0</td>
<td>10</td>
<td>.796</td>
<td>1</td>
<td>.681</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>14</td>
<td>.737</td>
<td>13</td>
<td>.636</td>
<td>13</td>
</tr>
<tr>
<td>2</td>
<td>19</td>
<td>.914</td>
<td>11</td>
<td>.714</td>
<td>11</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>.929</td>
<td>8</td>
<td>.786</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>.964</td>
<td>8</td>
<td>.804</td>
<td>10</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>.875</td>
<td>5</td>
<td>.871</td>
<td>7</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>.881</td>
<td>9</td>
<td>.905</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>.911</td>
<td>4</td>
<td>.964</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>.929</td>
<td>3</td>
<td>.929</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>3</td>
</tr>
</tbody>
</table>

45
Table 6
Summed Ranks for Pairs of Characteristics

<table>
<thead>
<tr>
<th>ALTERNATIVES</th>
<th>Second Best</th>
<th>Best</th>
<th>Third Best</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed-Pilot Sum of Ranks</td>
<td>12 10 12 12 9 12 12 11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance-Delay Sum of Ranks</td>
<td>11 9 11 11 8 11 11 10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact-Search Sum of Ranks</td>
<td>13 11 13 13 10 13 13 12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Armament-Radar Sum of Ranks</td>
<td>8 6 8 8 5 8 8 7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 7
Ranks for Single Characteristics

<table>
<thead>
<tr>
<th>ALTERNATIVES</th>
<th>Second Best</th>
<th>Best</th>
<th>Third Best</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed</td>
<td>6</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Pilot</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>Distance</td>
<td>3</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Delay</td>
<td>8</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Contact</td>
<td>8</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Search</td>
<td>5</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>Armament</td>
<td>7</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Radar</td>
<td>1</td>
<td>4</td>
<td>7</td>
</tr>
</tbody>
</table>
Table 8
Frequency of Choice of Each Alternative in Each Problem

<table>
<thead>
<tr>
<th>Alternative</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>n</th>
<th>Sum Over Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( f_{11} )</td>
<td>( f_{12} )</td>
<td>( f_{13} )</td>
<td>( f_{1n} )</td>
<td>( F_1 )</td>
</tr>
<tr>
<td>2</td>
<td>( f_{21} )</td>
<td>( f_{22} )</td>
<td>( f_{23} )</td>
<td>( f_{2n} )</td>
<td>( F_2 )</td>
</tr>
<tr>
<td>( k )</td>
<td>( f_{k1} )</td>
<td>( f_{k2} )</td>
<td>( f_{k3} )</td>
<td>( f_{kn} )</td>
<td>( F_k )</td>
</tr>
<tr>
<td>Sum Over Alternatives</td>
<td>( F_{-1} )</td>
<td>( F_{-2} )</td>
<td>( F_{-3} )</td>
<td>( F_{n} )</td>
<td>( N )</td>
</tr>
</tbody>
</table>
REFERENCES


NOTE:-

The data described in this paper were collected while the author was at
the U. S. Naval Research Laboratory, Washington, D. C. The assistance
of Miss Elizabeth C. Smith in collecting the data is gratefully acknowledged.
Electronic Systems Division, L. G. Hanscom Field, Mass.  
Rpt No. ESD-TDR-62-48. HUMAN DATA PROCESSING  
LIMITS IN DECISION MAKING. July 1962. 50 p., incl.  
figs., tables, 8 refs.  
Unclassified report

Four experiments are described in which subjects were  
required to choose among alternatives on the basis of two,  
four, six, or eight relevant facts. Both decision quality  
and decision time were measured. Presenting more than  
four facts caused a decrease in decision making efficiency.

1. Psychology  
2. Reasoning  
1. AFSC Project 2806  
Task 280603  
II. Hayes, J. R.  
III. In ASTIA collection

Electronic Systems Division, L. G. Hanscom Field, Mass.  
Rpt No. ESD-TDR-62-48. HUMAN DATA PROCESSING  
LIMITS IN DECISION MAKING. July 1962. 50 p., incl.  
figs., tables, 8 refs.  
Unclassified report

Four experiments are described in which subjects were  
required to choose among alternatives on the basis of two,  
four, six, or eight relevant facts. Both decision quality  
and decision time were measured. Presenting more than  
four facts caused a decrease in decision making efficiency.

1. Psychology  
2. Reasoning  
1. AFSC Project 2806  
Task 280603  
II. Hayes, J. R.  
III. In ASTIA collection
Four experiments are described in which subjects were required to choose among alternatives on the basis of two, four, six, or eight relevant facts. Both decision quality and decision time were measured. Presenting more than four facts caused a decrease in decision-making efficiency.