

STIMULUS INFORMATION AS A DETERMINANT OF REACTION TIME¹

RAY HYMAN

The Johns Hopkins University²

In the typical reaction-time experiment, *S*'s reaction time is greater when he has to respond differentially to one of two equally probable stimuli instead of to just one stimulus. In fact, Merkel (2), using one to ten alternatives, has demonstrated that when *S* has to respond to one stimulus chosen from a number of equally probable alternatives, his reaction time increases with the number of alternatives.

The fact that *S*'s response to stimulus *A* takes more time when *A* is one of several rather than one of two equally probable alternatives is of intrinsic interest. But it becomes even more significant when looked at from the standpoint of modern communication theory. In communication theory the amount of *information* which a message conveys is an increasing function of the number of possible messages from which that particular message could have been selected.³ The *S*'s reaction time seems to behave,

¹ Based upon a thesis submitted to the Faculty of Philosophy, The Johns Hopkins University in partial fulfillment of the requirements for the degree of Doctor of Philosophy. The author wishes to express his gratitude to his adviser, Dr. W. R. Garner, and to Drs. A. Chapanis and H. W. Hake for assistance in the preparation of this paper.

² This research was done in cooperation with the Systems Coordination Division, Naval Research Laboratory under Contract N5-ori-166, Task Order I, between the Office of Naval Research and The Johns Hopkins University. This is Report No. 166-I-163, Project Designation No. NR 507-470, under that contract.

³ The amount of information conveyed by the selection of a particular message is defined as the logarithm to the base 2 of the reciprocal of the a priori probability that this particular message will be selected.

under certain conditions, in a manner analogous to this definition of information. When a stimulus is chosen to which *S* must make a discriminatory response, his reaction time seems to be a monotonically increasing function of the number of possible stimuli from which the stimulus can be chosen.

Thus, the choice reaction-time experiment can be looked upon as a model of a communication system. The display represents a transmitter of information. Each alternative stimulus or *signal* represents a *message*; more information can be transmitted the greater the number of messages from which one can be chosen. The *channel* over which the signal is transmitted can be considered as the air space between the light and *S*, and might also include part of *S*'s visual afferent system. The *S* also acts as a *receiver* or decoder in that at some point he decodes the signal into its message and reacts with the appropriate response (the *destination* of the information).

Setting of the problem.—The experimental task involved varying the amount of information in the display and observing the corresponding changes in *S*'s reaction time to a stimulus presentation. The display was a matrix of lights, each light representing a message. The *S*'s reaction time was registered by means of a voice key and timer. The average amount of information accompanying the presentation of a single stimulus was varied by (a) varying the number of equally probable alternatives from which a choice could be made, (b)

altering the probability of occurrence of particular choices, and (c) introducing sequential dependencies between successive choices of alternatives.

These three ways of varying the amount of information per stimulus have been incorporated into a single formula by Shannon in the formulation of his mathematical theory of communication (3). Such a unified formula represents a convenience for the communication engineer. From the psychologist's viewpoint, however, the psychological equivalence of these three ways of varying information must be demonstrated empirically before Shannon's formula can be applied to the human component of the communication system. To demonstrate such an equivalence was one of the primary aims of the present study.

Statement of the problem.—The particular hypotheses investigated can be stated as follows:

1. Reaction time is a monotonically increasing function of the amount of information in the stimulus series.

2. The regression of reaction time upon amount of information is the same whether the amount of information per stimulus is varied by altering the number of equally probable alternatives, altering the relative frequency of occurrence of particular alternatives, or altering the sequential dependencies among occurrences of successive stimuli.

These hypotheses assume that (a) S's responses are completely determined by the stimulus series, and (b) the occurrences of the successive stimuli do not alter S's knowledge of the statistical properties of the stimulus series as a whole.

The first assumption demands a one-to-one correspondence between stimulus and response series. Hick's experiment (1) indicates that S can decrease his reaction time to a given

amount of information at the expense of an increased proportion of incorrect responses. The present study excludes this possibility by demanding an errorless performance on the part of S.

The second assumption excludes those situations wherein S gains new knowledge concerning the statistical structure of the stimulus series as the series progresses in time. The hypotheses assume that S's *average uncertainty per stimulus presentation* remains constant throughout a particular stimulus series. The present experimental design helps to produce such a situation by (a) never presenting S with the same series for a given condition more than once, (b) instructing S concerning the statistical properties of the series, and (c) giving S preliminary practice on a series of similar statistical construction.

PROCEDURE

Apparatus.—The apparatus consisted of a visual display, a voice key, a chronoscope, and a control panel. The display was a square matrix of 36 small lights. These 36 lights were arranged in six rows and six columns forming a square 3 in. on the side. The matrix was so placed that the length of a side of the display made a visual angle of 5° from where S sat. The diameter of each light in the display made a visual angle of approximately 36'.

Only eight of the lights were used: the four lights which made up the corners of the outer square of lights and the four corners of the next inner square of lights. Such a grouping seemed to minimize the confusions of one light with another on the matrix by S. The S designated each light by the distinctive names: *Bun, Boo, Bee, Bore, By, Bix, Bev, and Bate*, respectively.

A masonite screen separated E, the control panel, and the chronoscope from S. The E preselected a light by means of a rotary switch. After giving S a warning signal, he closed a switch (2 sec. later) which simultaneously turned on the light and started the clock. The S's vocal response stopped the clock by means of a throat microphone which activated an electronic voice key.

Subjects.—The four Ss were male undergraduates, ages 18–22. Each S attended more than

40 experimental sessions over a 3-mo. period; approximately 15,000 reaction times were recorded for each S. One of the Ss, F. K., also took part in 11 sessions of the pilot study during the month previous to the 3 mo. he served as S in the actual experiment.

Experimental conditions.—In all there were three experiments, each making use of one of the three different methods of altering stimulus information. Each of these experiments consisted of eight specific conditions containing different amounts of stimulus information. In addition, five series of stimuli, each constructed separately but according to the same rules, were used for each experimental condition.

Experiment I.—The first experiment had eight conditions which differed from each other only in terms of the number of equally probable alternatives from which the stimulus could be chosen. The eight conditions involved numbers of alternatives ranging from one through eight; the respective bits of information attached to each condition were 0.00, 1.00, 1.58, 2.00, 2.32, 2.58, 2.81, and 3.00. This variable is essentially the same as used in the experiments of Merkel (2) and Hick (1).

Experiment II.—The second experiment had eight conditions which involved different numbers of alternatives and different probabilities of

the occurrence of these alternatives. Because of the greater complexity of these conditions, they are shown in Table 1. In Cond. 1, for example, there are just two alternatives, but one occurs nine times as often as the other. The stimulus information in bits is shown at the right for each component separately and for the weighted average of the total condition. In Cond. 3 there is a total of four different alternatives, one of which occurs 13/16 of the time, with the other three each occurring 1/16 of the time. Conditions 7 and 8 each involve the greatest number of different alternatives (eight), with the proportions of occurrences as shown.

Experiment III.—In the third experiment different numbers of alternatives were used, and in every condition all alternatives occurred equally often. But the probability of the occurrence of a particular alternative depended upon the immediately preceding alternative in a manner shown in Table 2. Condition 1, for example, consisted of two alternatives. Whenever one of these alternatives occurred, it would be followed by the other alternative 8/10 of the time and by itself 2/10 of the time. In Cond. 3, with four alternatives, the occurrence of a stimulus meant that the probability of its recurring on the next trial was 7/10; the probability of some other alternative occurring was 1/10. Conditions 6, 7, and 8 were constructed in the same manner as conditions in Exp. I with the exception that a particular stimulus was never followed by itself.

The experimental sessions.—An experimental session lasted 1 hr. Each S attended one session a day, five days a week, at the same hour each day. Three series (from three different conditions) were run during each session. Since each series consisted of 120–128 stimuli, each session yielded a total of 360–384 reaction times. For experimental convenience each series was split into two parts of 60–64 stimuli each. By designating the three series as A, B, and C, the typical counterbalanced order of any session can be symbolized in this manner: Practice on A, A, Practice on B, B, Practice on C, C, 5-min. rest, C, B, A.

RESULTS

Experiment I.—When the amount of information per stimulus presentation was varied by varying the number of equally probable stimulus alternatives from which *E* could choose the stimulus, the linear correlation between amount of information and reaction time was .991, .982, .980, and .979 for each of the four Ss. There

TABLE 1

THE EIGHT CONDITIONS FOR EXPERIMENT II
AND THE CORRESPONDING AMOUNTS OF
INFORMATION IN BITS PER STIMULUS
PRESENTATION

Cond.	Number of Alternatives	Probability of Occurrence	Log ₂ 1/p	Av. Amount of Information in Cond.
1	1	9/10	0.15	0.47
	1	1/10	3.32	
2	1	8/10	0.32	0.72
	1	2/10	2.32	
3	1	13/16	0.30	0.99
	3	1/16	4.00	
4	1	15/20	0.42	1.39
	5	1/20	4.32	
5	1	4/8	1.00	1.75
	1	2/8	2.00	
	2	1/8	3.00	
6	1	5/10	1.00	2.16
	5	1/10	3.32	
7	1	8/16	1.00	2.38
	1	2/16	3.00	
8	6	1/16	4.00	2.75
	2	4/16	2.00	
	2	2/16	3.00	
	4	1/16	4.00	

TABLE 2

THE EIGHT CONDITIONS FOR EXPERIMENT III AND THE CORRESPONDING AMOUNTS OF INFORMATION IN BITS PER STIMULUS PRESENTATION

Cond.	Number of Lights $p_a(x)$	Probability (p) of Stimulus x Occurring Following Stimulus a	$\text{Log}_2 1/p$	Average Amount of Information in Cond.
1	2	$p_a(b) = 8/10$ $p_a(a) = 2/10$	0.32 2.32	0.72
2	3	$p_a(a) = p_a(c) = 1/10$ $p_a(b) = 8/10$	3.32 0.32	0.92
3	4	$p_a(a) = 7/10$ $p_a(b) = p_a(c) = p_a(d) = 1/10$	0.51 3.32	1.36
4	4	$p_a(b) = 3/6$ $p_a(a) = p_a(c) = p_a(d) = 1/6$	1.00 2.58	1.79
5	8	$p_a(b) = 9/16$ $p_a(a) = p_a(c) = \dots p_a(h) = 1/16$	0.82 4.00	2.21
6	3	$p_a(a) = 0$ $p_a(b) = p_a(c) = 1/2$	0.00 1.00	1.00
7	5	$p_a(a) = 0$ $p_a(b) = p_a(c) = \dots p_a(e) = 1/4$	0.00 2.00	2.00
8	8	$p_a(a) = 0$ $p_a(b) = p_a(c) = \dots p_a(h) = 1/7$	0.00 2.81	2.81

was no systematic tendency for the variances of reaction times within each condition to be correlated with the means of the conditions.

Three Ss showed a slight, but statistically significant, practice effect from series to series. For G. C. and F. P. this practice effect accounted for only 3% of the total variance among series. For L. S., however, this practice effect amounted to 11.5% of the total variance among his 40 series. This fact will be significant in later interpretation of some results concerning this S's regression lines.

As was to be expected, F. K., who had a month's preliminary training on Exp. I as part of the pilot study, showed no practice effect from series to series.

It is interesting to note that Merkel (2) and later Hick (1) also obtained reaction time as a function of the number of stimulus alternatives. Both experimenters used finger keys rather than voice keys. When Merkel's data are converted into reaction time as a function of the amount of information in the stimulus display, the linear

regression accounts for 99% of the variance among his reaction-time means. Hick's graphed points indicate that the linear correlation between reaction and information is approximately of the same magnitude as in the present study.

Experiment II.—The linear correlation between reaction time and information when the amount of information was varied by varying the relative frequency of occurrence of the various stimuli with respect to each other was .989, .965, .994, and .952 for each of the four Ss. As in Exp. I, the variances for each condition were not systematically related to the means. No practice effect was evident.

Experiment III.—When the amount of information was varied by introducing sequential dependencies among the successive stimulus presentations, the linear correlation between reaction time and information for each of the four Ss was .972, .965, .874, and .934, respectively. These correlations are systematically lower than those for the previous two experiments. In all four

cases the nonlinear variance is significant at the 1% level.

The first five conditions in Table 2 produced reaction times which were significantly lower than the reaction times for the latter three conditions

after the means were adjusted for the linear trend. Practically all of the nonlinear variance for each of the four Ss was produced by the one degree of freedom used to compare these two groups of conditions; the nonlinear

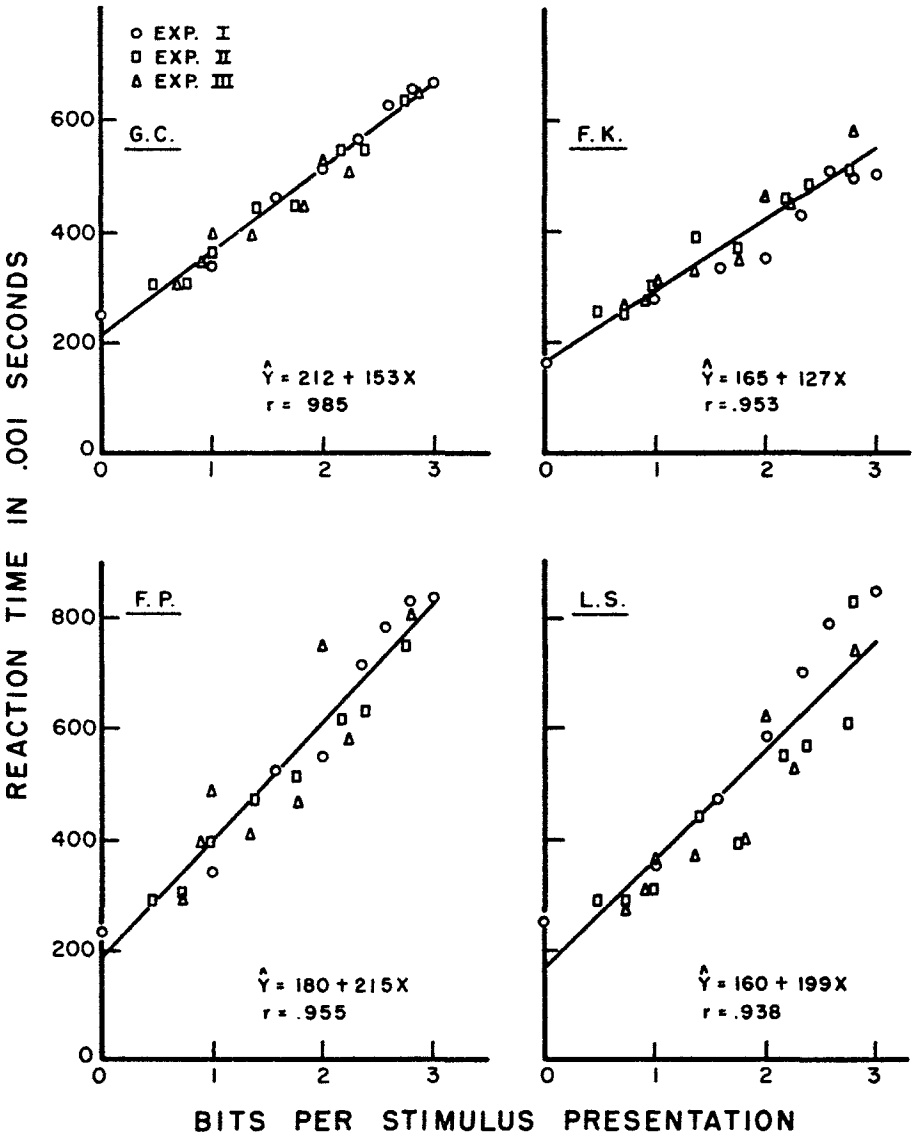


FIG. 1. Reaction time as a function of stimulus information (expressed in bits) when amount of information is varied in three different ways. Data are for Ss G. C., F. K., F. P., and L. S., respectively.

variance associated with the remaining five degrees of freedom was non-significant.

The high reaction times for the three conditions wherein no immediate repetition of a stimulus was allowed raise some interesting questions. On the basis of information theory, the restriction "no repetition" lowers the amount of information per stimulus presentation since it effectively decreases the number of alternatives by one. Instead of lowering *S*'s reaction time, however, this restriction caused his reaction time to be higher than if no such restriction had been imposed.

Comparison of the three experiments.—The second experimental hypothesis stated that the three experiments would produce identical regression lines. In the case of three *S*s, the null hypothesis that the three regression lines do in fact coincide could not be rejected at the 5% level of confidence. For L. S., however, the null hypothesis had to be rejected at the 1% level of confidence. This lack of coincidence could not be attributed to a discrepancy among the three regression coefficients; the difference was due to a systematic displacement of the regression line for Exp. I upwards on the ordinate relative to the lines for Exp. II and III.

This systematic displacement among the regression lines of L. S. can be attributed to the relatively large practice effect which this *S* displayed on Exp. I. Most of the series for Exp. I were run during the first six weeks of the experimental period, whereas most of the series for Exp. II and III were run during the latter half of the experimental period. L. S., it will be recalled, showed a practice effect which accounted for 11.5% of the variance among his series for Exp. I, whereas he presumably had reached a plateau by the time he participated in most of the

series of Exp. II and III (wherein he showed no practice effect). Consequently his regression line for Exp. I was displaced slightly upwards on the ordinate.

The joint regression line.—The joint regression line fitted to the combined points of the three experiments produced a linear correlation between reaction time and amount of information for each of the four *S*s of .985, .953, .955, and .938, respectively (see Fig. 1). In other words, when we vary information in three distinct ways over a limited range of from 0.00 to 3.00 bits, a linear regression of reaction time on information accounts for as much as 97.0, 91.1, 91.1, and 88.8% of the total variance among mean reaction times for the 24 experimental conditions for each of the four *S*s.

DISCUSSION

So far this paper has presented empirical relationships which suggest that reaction time can be considered a linear function of stimulus information within the range of 0.00 to 3.00 bits. More important than the shape of this function are the factors which operate to bring it about. This particular study was set up to discover the type of relationship which exists rather than to delve into the causes of this relationship. Nevertheless, an examination of the data suggests some factors which may partially account for the resultant function.

Nonadditive combination of components within conditions.—The conditions for Exp. II and III can be considered as made up of two or more components, each component having a different amount of information associated with its presentation (see Tables 1 and 2). Condition 1 in Table 1, for example, has two components, one with 0.15 bits and the

other with 3.32 bits. On the hypothesis that reaction time behaves in a manner analogous to the measure of information, we would expect the mean reaction time to each of the components within a condition to fall on the regression line which was fitted to the over-all means of the conditions. To illustrate, one might expect a constant mean reaction time to an event which occurs with probability $\frac{1}{2}$ regardless of how many other elements there are in the series. But such was not the case. In both Exp. II and III the components within a condition interacted with each other in such a way that the reaction time to the low information component was higher, and the reaction time to the high information component was markedly lower, than would be predicted on the basis of the regression line fitted to the means of the conditions. This effect was very marked and occurred without a single exception for all four Ss in both Exp. II and III.

An example will illustrate the nature of this effect. Take the condition in Table 1 (Cond. 3) which has 0.99 bits associated with it. For G. C., the regression line fitted to the reaction times of the 24 experimental conditions would predict a mean reaction time for this condition of 363 msec. The observed mean reaction time for this condition was 361 msec., a figure which agrees closely with the expected value. This condition consisted of two components, one with 0.30 bits of information and the other with 4.00 bits of information attached to its occurrence. On the basis of the over-all regression line to G. C.'s points we would predict mean reaction times for these two conditions of 258 and 824 msec., respectively. If we weight each of these expected values by the probability of occurrence of the component with which it is attached,

13/16 and 3/16, respectively, our weighted predicted average comes out to be 363 msec. The observed mean reaction time for the component with 0.30 bits was 306 msec., which was 48 msec. higher than the predicted value, and the observed mean reaction time to the component with 4.00 bits was 585 msec. or 235 msec. lower than the predicted value. Yet, when we weight these component means by their frequency of occurrence, we come up with an observed mean of 361 msec., agreeing very closely with the predicted mean. Such a situation holds for all the conditions and their components in Exp. II and III and for all four Ss.

These analyses mean that we cannot predict, on the basis of the regression line fitted to the means of the conditions, what the mean reaction times will be to the components which make up a condition. If, however, we know what the components of the condition are, we can predict what the combined mean for this condition will be on the basis of the over-all regression. The components, of course, do not combine additively. But they interact in such a manner that the condition means behave *as if* the components combined additively. If we are interested only in the behavior of the condition means, the assumption of additive combination of the components will serve our purposes. If, however, we are interested in the behavior of the components making up the conditions, we must find different laws and equations.

Reaction time as a function of the number of other stimuli intervening between successive occurrences of a stimulus.—Another interesting factor determining the reaction time was first observed during the pilot study. The *E* observed that whenever a stimulus was immediately followed by

itself in the series, *S* seemed to respond unusually fast to it. This effect was apparently independent of *S*'s verbalized expectancy of the stimulus for he reported that in such cases he was not expecting the stimulus to follow itself. An examination of the data showed that this phenomenon was quite marked for the situation with four or more alternatives and steadily declined until it disappeared or became slightly negative for the case with just two alternatives.

As can be seen in Fig. 2, the experimental data from the four *S*s support the observations made from the pilot study. In this figure reaction time is plotted as a function of the number of stimuli intervening between successive occurrences of a particular stim-

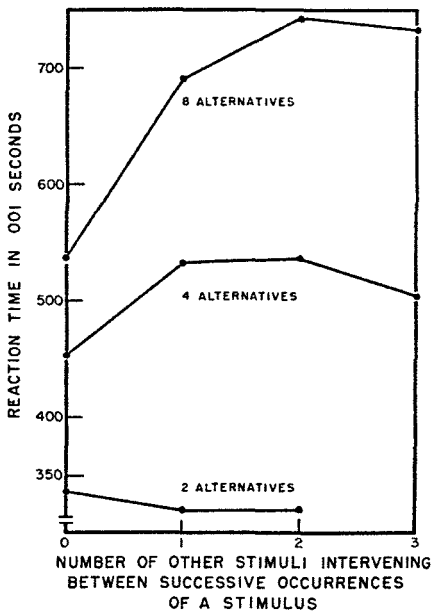


FIG. 2. Reaction time as a function of the number of other stimuli intervening between successive occurrences of a particular stimulus within a series. Each point represents an average of the data from four *S*s. The three parameters are the conditions with eight, four, and two alternatives from Exp. I.

ulus in the series. The three parameters plotted are the conditions with two, four, and eight alternatives from Exp. I. Although the data are averaged over the four *S*s, the individual curves differed from each other only in displacement along the ordinate.

For more than two alternatives the general trend of the functions seems to be parabolic; the function seems to reach a maximum at around a displacement of one or two stimulus presentations and then come down again. The effect is more marked the greater the number of alternatives. Presumably two factors are operating to produce this bow-shaped curve. One is *S*'s verbalized introspection that a stimulus which has not appeared for some time in the series is reacted to more quickly than ordinarily because of the greater expectancy now attached to its appearance. This "verbal expectancy" apparently accounts for the fact that the reaction times begin to get lower than the maximum for large displacements. The second factor seems to consist of some sort of residual effect produced by just having seen and reacted to a particular stimulus; this effect seems to facilitate reaction to this stimulus if it reappears within a finite time interval. For eight alternatives this facilitation seems to last for at least a displacement of one stimulus presentation. For two alternatives this facilitation does not affect the function; perhaps it is at its maximum throughout the series and therefore does not show in the function for two alternatives.

Evidence for periodicity.—The data suggest a third factor which may have to be taken into account in the final story concerning the role of stimulus information as a determinant of reaction time. Separate frequency distributions were made of the reaction

times in each condition. For the situations wherein we have two or more alternatives the distributions tended to be multimodal. In almost every distribution a peak, clearly differentiated from the remainder of the distribution, regularly occurred at approximately *S*'s simple reaction time. The remainder of the distribution in each case, however, did not present such a consistently clear picture. The variability of the data was such that it is impossible to make any definite conclusions about the number of peaks and their location on the abscissa. Furthermore, if there is any natural periodicity in the distribution it seems to have been confounded with a periodicity suggested by the grouping of stimuli on the display. Since the distributions seem to support several possible and different models as to what is happening, further discussion must be postponed until more data are available.

SUMMARY

The reaction time to a visual stimulus was investigated as a function of the amount of information conveyed by that stimulus. The amount of information in the stimulus was varied by varying (a) the number of equally

probable alternatives from which it could be chosen, (b) the proportion of times it could occur relative to the other possible alternatives, and (c) the probability of its occurrence as a function of the immediately preceding stimulus presentation.

The reaction time to the amount of information in the stimulus produced a linear regression for each of the three ways in which information was varied.

The three regression lines obtained by the three separate ways of varying the amount of information were found to coincide for three *S*s. In the case of the fourth *S*, a systematic displacement of one of his regression lines was attributed to a relatively large practice effect which he showed for that experiment.

(Received for early publication
November 20, 1952)

REFERENCES

1. HICK, W. E. On the rate of gain of information. *Quart. J. exp. Psychol.*, 1952, 4, 11-26.
2. MERKEL, J. Die zeitlichen Verhältnisse der Willensthatigkeit. *Philos. St.*, 1885, 2, 73-127.
3. SHANNON, C. E., & WEAVER, W. *The mathematical theory of communication*. Urbana: Univer. of Illinois Press, 1949.