INFORMATION CONTENT AS A DETERMINANT OF AVERAGE REACTION TIMES TO PURE TONES
A Preliminary Investigation

by Arthur Albert

Submitted in partial fulfillment of the requirements

for the degree of Bachelor of Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
(1956)

Signature of Author Signature redacted

Mathematics May 21, 1956
Dept. date

Certified by Signature redacted
Thesis Advisors

ACKNOWLEDGMENTS

I wish to express my thanks to my thesis advisors, Professors Walter Rosenblith and George Thomas. In addition, a large measure of gratitude must be extended to Robert Brown, who spent a good deal of time solving the instrumentation problems associated with this investigation, and to Dr. Thomas Sandel who made many helpful suggestions concerning the experimental procedure.

Last, but not least by any means, I wish to thank my subjects, Raylene Burdman, Aurice Alberts, Kathie Manos and Dan Geisler, who devoted their time willingly, and without whose help this experiment would not have been possible.

A.A.

ABSTRACT

The reaction time to an auditory stimulus (pure tone) was investigated as a function of the average amount of information conveyed by that stimulus. The amount of information in the stimulus was varied by varying a) the number of alternatives and b) the proportion of times any one of the stimuli could occur, once the number of alternatives had been fixed.

The experiment was in three parts. In the first, simple reaction times were measured for all the stimuli that would be involved in the subsequent stages of the experiment. In the second stage, the number of different stimuli was fixed at two and the average information per stimulus was varied from zero to one bit. In the third stage, the number of different stimuli was fixed at four and the average information per stimulus was varied from zero to two bits.

The regression lines found for the last two parts ($\bar{t}$ on $\bar{I}$) were linear for all S's. Furthermore, for three of the four S's, the two lines coincided.

TABLE OF CONTENTS

Introduction

In a typical reaction time experiment, the subject's reaction time is greater when he must discriminate between two equally probable stimuli instead of simply responding to a single stimulus. Merkel[1], using one to ten alternatives, has demonstrated that when S has to respond correctly 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 reaction time to a stimulus A is greater when A is one of several, rather than one of two equally probable stimuli is quite interesting in itself. When viewed from the standpoint of information theory, the phenomenon takes on new scope and in the process, many new questions are raised.

The reaction time experimental situation may be viewed as a process of transmitting information. A source, transmits a signal over a communications channel. The signal reaches its destination (the subject) and S must act as a receiver or decoder. Upon completion of the decoding process, he must display the correct response and thus transmit the message to the experimenter.

In this experiment we will concern ourselves mainly with the following parameters: If there are k possible stimuli, and the subject knows that the j-th stimulus occurs with probability p_j , then we define the information contained in that stimulus as $I_j = -\log_2 p_j$. If a message consists of a sequence of such stimuli, the average information contained in each stimulus is $\bar{I} = -\sum_{j=1}^{k} p_j I_j = -\sum_{j=1}^{k} p_j \log_2 p_j$ where I and $\bar{I}$ are measured in bits.

The other parameter which will occupy a central position in the following experiment is, of course $\bar{t}$, the average reaction time associated with the previous sequence of stimuli.

We shall conduct an exploratory investigation which concerns itself with the following question: What can be said about the dependence of decision reaction times upon the average information "contained in" the stimulus to which S is responding?

Procedure

general:

S was placed in an anechoic room and was given a pair of earphones wired for monaural listening. On the arm of his chair was mounted a box. On the top surface of the box was a square array of four circular metallic plates about the size of pennies. The center-to-center distance between the plates was approximately three cm. At the center of the square was a "rest" position (see diagram 1). The box was so designed, that the plates were raised to 45 volts with respect to S. S wore an electrode on his fourth finger so that his hand would be at ground potential. The onset of the stimulus (a pure tone) tripped a decimal counter. When the subject responded by touching the correct plate, a negative pulse was created. This pulse, after passing through an amplifier, stopped the counter.

apparatus:

A pulse, generated by a pulse generator, triggered a fixed delay circuit. A fixed interval after the delay circuit had been activated, the warning light flashed on for one second. A second pulse left the fixed delay at this time and passed through a variable delay which controlled the latency between the warning and stimulus. After the pulse left the variable delay, it activated an interval timer which immediately sent out two signals: one

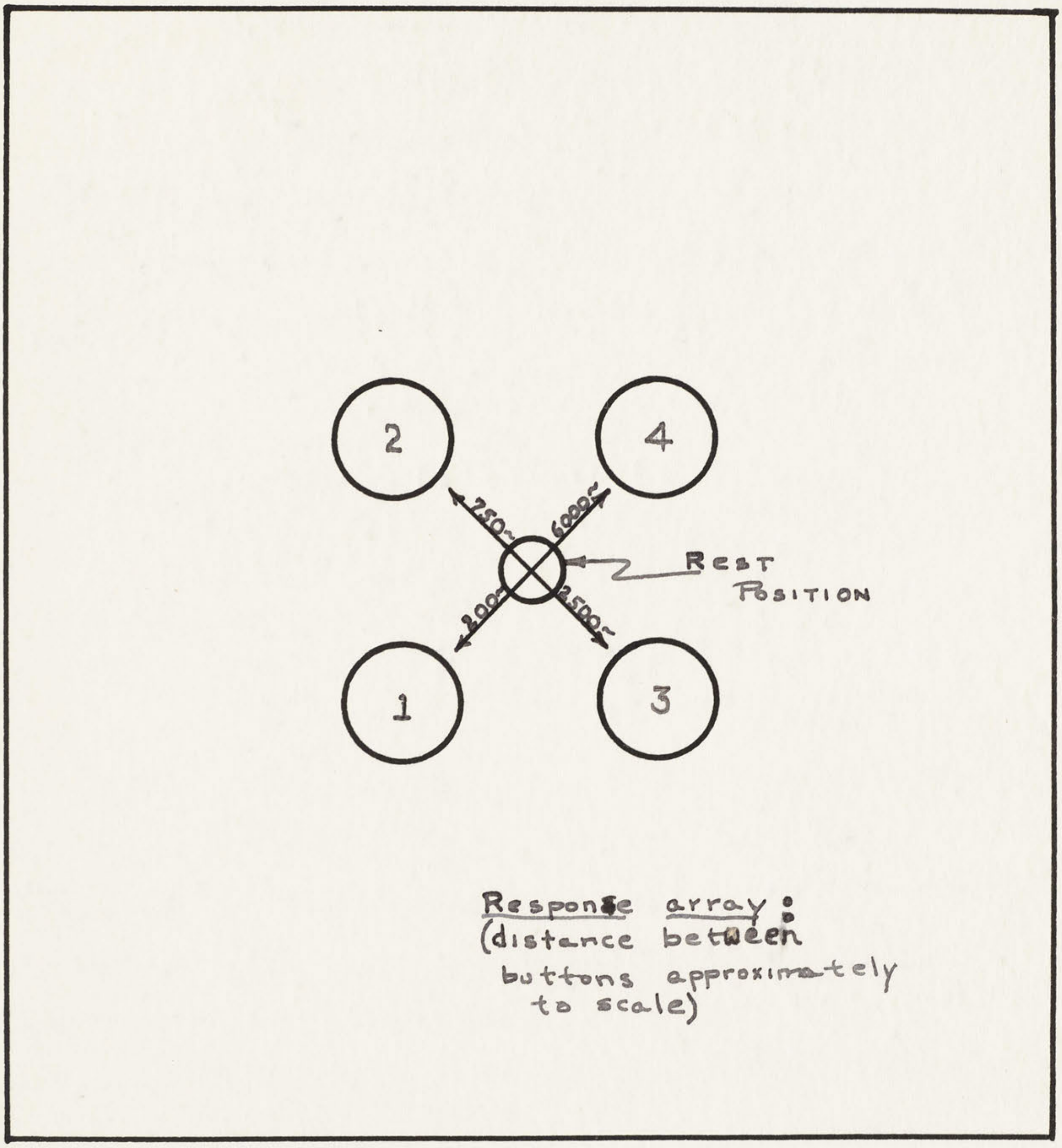

2
4
750~
6000~
Rest Position
800~
2500~
1
3
Response array:
(distance between buttons approximately to scale)

DIAGRAM 1.

to start the decimal counter, the other to activate an electronic switch. Two hundred milliseconds later, another pulse from the interval timer closed the electronic switch. In the meantime, the electronic switch had passed a sinusoidal wave (with a rise-fall time of 10 milliseconds) through a series of attenuators and amplifiers to the subject's earphones. (See diagram 2)

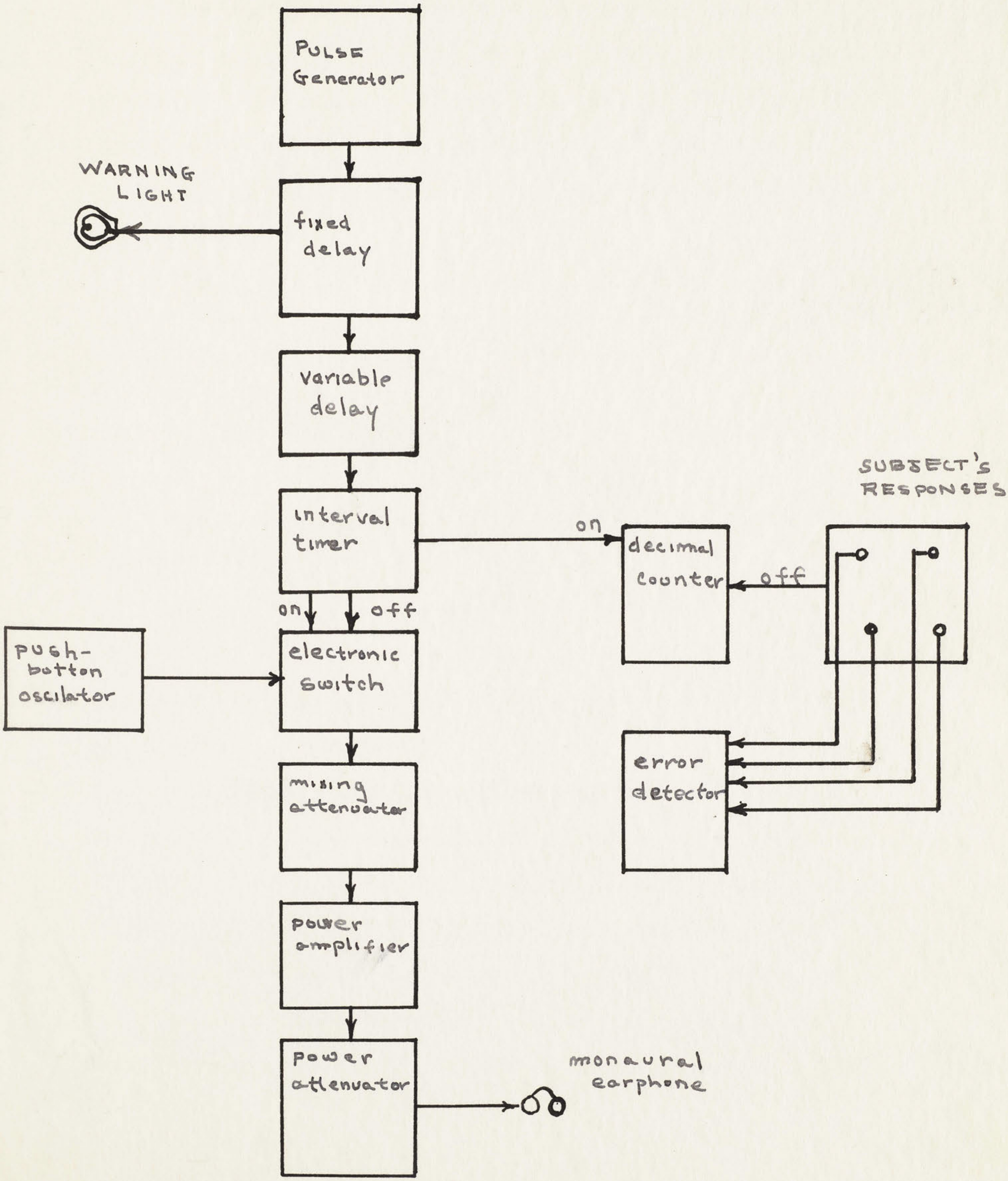
The subject responded by touching the correct plate. The pulse created by this response stopped the decimal counter and the reading on the counter indicated the desired latency in milliseconds. If S made an incorrect response, this was recorded by an error detector which flashed a light on the experimenter's instrument panel. All stimuli were at the same intensity: -20 db. re 1 volt rms.

subjects:

Four subjects aged 18-25 were used. Three were female, one was male. All had normal hearing. Each subject was trained at one frequency as follows: Sitting in the chamber, the subject would see a warning light flash. One half to four seconds later, the subject would hear a 2500 cycle tone (-20 db. intensity). He was instructed to move his finger from the rest position to plate #3, then return to the rest position and await the next warning-stimulus. This training continued until some degree of stability was achieved in the subject's variability. (For all subjects four runs of fifty stimuli seemed sufficient.)

The next step was to determine simple reaction times for each of the four frequencies which were to be used as stimuli in subsequent stages of the experiment.

DIAGRAM 2: Block Diagram of experimental apparatus



experiment I:

Four runs of 50 stimuli apiece were taken. S was instructed to put his finger on the rest position, await the warning light, touch the proper plate when the stimulus occured, then return to the rest position. All runs were taken at an intensity of -20 db. re 1 volt rms. The experimental conditions are best summarized by the following table:

run #	stimulus frequency (cps)	proper response (plate #)
1	200	1
2	750	2
3	2500	3
4	6000	4

experiment II:

Four runs of 50 stimuli apiece were taken. S was instructed in the following manner: "In this run, you will be presented with a sequence of 50 tones. Some of them are tone "a" (a sample of tone "a" was given here), some will be tone "b" (a sample of tone "b" was given here). These tones will be presented in a random order. However, the probability of tone "a" occuring is .5; the probability of tone "b" occuring is .5. This does not necessarily mean that exactly half are "a"'s and half are "b"'s.

"When you hear tone "a", you are to touch plate "a". When you hear tone "b", you are to touch plate "b"." (Since tones "a" and "b" were chosen from the four tones in the preceding experiment, the plates which S was instructed to touch, corresponded to the plates he had touched (for the tones in question) in experiment I.)

After a thirty minute practice period, the subjects were assumed to be trained and the experiment commenced. In this stage of the experiment, the warning period was fixed at two seconds. (Since a choice situation was involved, the subject would have no way of anticipating the stimulus by watching the warning light and "jumping the gun".)

This experiment was repeated three times (to make a total of four runs in all) for each S. Each time, the average information input was altered over a range from 0 to 1 bit by varying the probabilities and telling the subject what the respective probabilities were to be. In all cases the probabilities were such that $p_b \geq p_a$ (subject, of course to $p_a + p_b = 1$) and if the frequencies involved were f_a and f_b , then $f_b > f_a$. These conditions are best summarized by the following table and table 3:

subject	f_a (cps)	f_b (cps)
RB	200	2500
AA	200	6000
KM	200	750
DG	750	2500

experiment III:

The subjects were given the following instructions: "You will be presented with a sequence of 50 tones. Some will be at a frequency of 200 cps (a sample was given here), some will be at a frequency of 750 cps (a sample was given here),
" " " " " " " 2500 " " " " " "
" " " " " " " 6000 " " " " " ".

"The tones will be presented in random order. However, the probabilities for each tone are as follows: Calling the tones 1, 2, 3, and 4 respectively, $p_1 = p_2 = p_3 = p_4 = \frac{1}{4}$. This does not mean that exactly 25% will be 1's, 25% will be 2's etc.

"When you hear tone 1, touch plate #1; when you hear tone 2, touch plate #2; etc. "

The experiment was repeated four times, the average information being varied by varying the probabilities in a manner best described by the following table (also see table 4):

run	p_1	p_2	p_3	p_4
1	.25	.25	.25	.25
2	.10	.30	.40	.20
3	.08	.14	.60	.18
4	.04	.12	.80	.04
5	.02	.06	.90	.02
6*	.00	.00	1.00	.00

* This point was taken as the simple RT for tone 3 above.

In this instance, corresponding runs were identical for all subjects. Since the probability sets were fairly complicated, a sheet was placed in view of the subject, upon which the pertinent probabilities were described.

For this experiment, subjects were, in general, slow to catch on, and it took 30 minutes to an hours' training before the subjects had mastered the technique of differentiating between the respective tones.

The experimenter was able to make note of all mistakes as they occured in both the second and third experiments.

The experimental sessions took place over a period of two months. Each session was divided into two to four sittings, each sitting lasting nine minutes. Each entry in tables 1, 5 and 6 represents the average of 50 points. In all, each subject's responses were measured for 650 stimuli.

Results

experiment I:

It is observed that all subjects display a tendency to have lower average reaction times to the 200 cps and 2500 cps stimuli than to the 750 cps and 6000 cps ones (see table 1 and fig. 1). This bias is probably due to the spatial relationship that exists between the plates. The correct responses to the 750 cps and 6000 cps tones were plates 2 & 4 respectively (see diagram 2). When the subjects were asked if they found it easier to go to any particular plate, it was unanimously agreed that it was "easier" to react in a downward direction (to plates 1 & 3) than in an upward direction. At first, this might appear to be an unwanted variable; ideally, the RT's would be such that a statistical test would admit the hypotheses that all came from the same population. But, in fact, this variation does not disturb the experiment at all. We are more interested in changes in responses and not so much in the responses themselves per-se. The statement will become clearer as the discussion progresses.

experiment II:

Each subject was presented with sets consisting of a different pair of tones (taken from the six possible combinations of four tones, taken two at a time). This was done so that any bias connected with a particular stimulus (which seemed to be uniformly present in all subjects) would be distributed over all subjects.

The plots are explained in the following manner: In fig.2, the abcissa is the average information contained in each stimulus. Average information is defined in the Shannon-sense: $\bar{I} = -(p_a \log_2 p_a + p_b \log_2 p_b)$. The ordinate is the average reaction time observed at a fixed average information input. The solid, straight lines represent the lines of linear regression of $\bar{t}$ on $\bar{I}$. The correlation coefficient between $\bar{t}$ and $\bar{I}$ for each S, is expressed in the following table:

subject	RB	AA	KM	DG
r	.99	.93	.98	.94

It is seen that a high degree of linearity does exist. Errors are tabulated below. The general entry in the table is the number of errors commited by a given S for a given run.

		average information input (bits)				
		1.000	.971	.722	.470	.000
subject	RB	0	0	1	0	0
	AA	2	1	2	0	0
	KM	0	0	0	0	0
	DG	0	0	0	0	0

Examination of fig. 3 suggests that, although there is no systematic trend for all S's, the variability of RT's in the two choice experiment are, in general, higher than those of the simple RT experiment.

experiment III:

In this experiment, all subjects listened to the same sets of stimuli for corresponding runs. Each run consisted of 50 stimuli, some of which were tone 1, some tone 2, etc. If the probability of stimulus number i occuring is p_i, then table 4 indicates the manner in which the sets (p_i) were changed from run to run. Fig. 4 shows the linear relationship between average RT and average information input where $\bar{I} = -\sum_{i=1}^{4} p_i \log_2 p_i$. The encircled points on fig.4 apply to comments in the summary section. The correlation coefficients between $\bar{t}$ and $\bar{I}$ are presented in the following table:

subject	RB	AA	KM	DG
r	.69	.76	.65	.72

Errors are tabulated below. The general entry in the table is the number of errors commited by a given S for a given run.

average information input (bits)

subject	2.00	1.84	1.58	1.01	0.60	0.00
RB	1	1	0	0	0	0
AA	4	2	0	2	0	0
KM	1	1	1	1	0	0
DG	4	3	2	2	1	0

Examination of fig.5 again reveals no systematic trend. However, it should be noted that the variability for the 4-choice experimental condition is almost everywhere greater than the variabilities associated with the 2-choice experiment.

Table 1:

MEAN REACTION TIMES OF EACH SUBJECT FOR EACH OF THE FOUR FREQUENCIES TAKEN SEPERATELY
(time measured in milliseconds)

	frequency in cps			
	200	750	2500	6000
RB	279	280	243	266
AA	197	214	207	213
KM	297	326	296	318
DG	195	235	178	238

Table 2:

STANDARD DEVIATIONS ASSOCIATED WITH EACH MEAN IN THE TABLE ABOVE

	Frequency in cps			
	200	750	2500	6000
RB	36	59	34	42
AA	29	17	15	13
KM	30	34	40	74
DG	27	15	14	24

Table 3:

AVERAGE INFORMATION CONTAINED IN A STIMULUS WHEN THERE ARE TWO STIMULI, a&b, WHICH OCCUR WITH PROBABILITIES p_a & p_b $(p_a + p_b = 1)$

p_a	.5	.4	.2	.1	0
I (BITS)	1.000	.971	.722	.470	.000

Table 4:

PROBABILITY COMBINATIONS OF THE STIMULI AND AVERAGE INFORMATION CONTAINED IN EACH STIMULUS FOR THE FOUR CHOICE SITUATION

(p_1, p_2, p_3, p_4)	I	
(0 , 0, 1, 0)	0	
(.02, .06, .90, .02)	.60	
(.04, .12, .80, .04)	1.01	(BITS)
(.08, .14, .60, .18)	1.58	
(.10, .30, .40, .20)	1.84	
(.25, .25, .25, .25)	2.00	

Table 5:

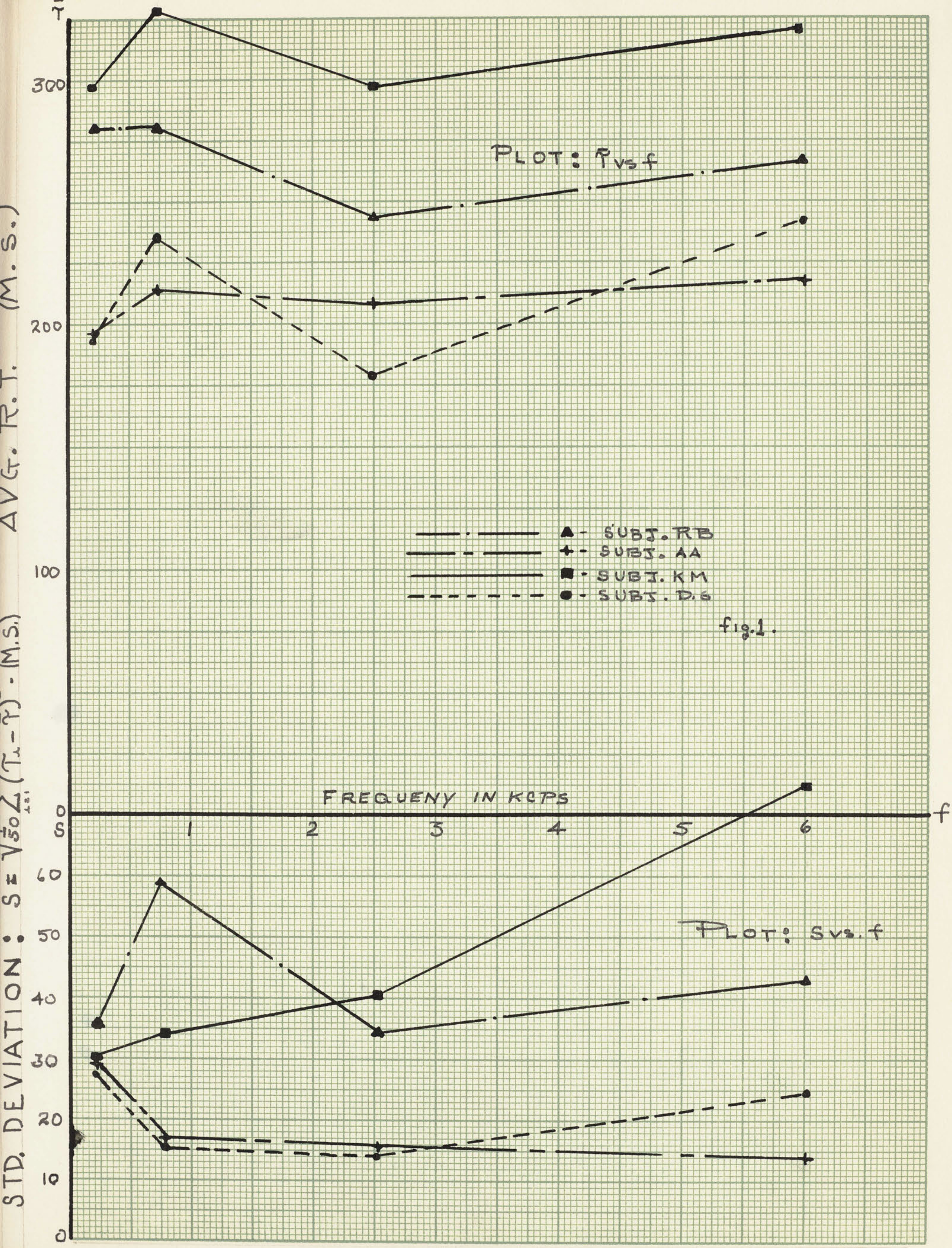
INFORMATION INPUT, AVERAGE REACTION TIME AND STD. DEVIATION FOR EACH SUBJECT IN THE TWO CHOICE SITUATION (Given in M.S.)

I (BITS)	($\bar{T}$; s) RB	AA	KM	DG
.000	(243;34)	(213;13)	(326;34)	(178;14)
.470	(363;82)	(264;71)	(354;48)	(211;70)
.722	(407;90)	(311;64)	(365;57)	(253;54)
.971	(480;93)	(340;55)	(373;64)	(272;73)
1.000	(455;103)	(376;87)	(360;55)	(320;75)

Table 6:

INFORMATION INPUT, AVERAGE REACTION TIME AND STD. DEVIATION FOR EACH SUBJECT IN THE FOUR CHOICE SITUATION (Given in M.S.)

I (BITS)	($\bar{T}$; s) RB	AA	KM	DG
.000	(243;34)	(213;13)	(326;34)	(178;14)
.60	(421;200)	(353;179)	(455;13)	(267;109)
1.01	(535;82)	(425;231)	(518;99)	(326;99)
1.58	(780;202)	(539;283)	(618;115)	(398;99)
1.84	(904;266)	(555;227)	(401;96)	(585;91)
2.00	(700;137)	(648;250)	(659;140)	(444;146)



PLOT: τ̄ vs f
PLOT: S vs. f
SUBJ. RB
SUBJ. AA
SUBJ. KM
SUBJ. D.G
FREQUENY IN KCPS
AVG. R.T. (M.S.)
STD. DEVIATION: $S = \sqrt{\frac{1}{50}\sum_{i=1}^{50}(T_i - \bar{T})^2}$ · (M.S.)
300
200
100
0
60
50
40
30
20
10
1
2
3
4
5
6
f

fig.1.

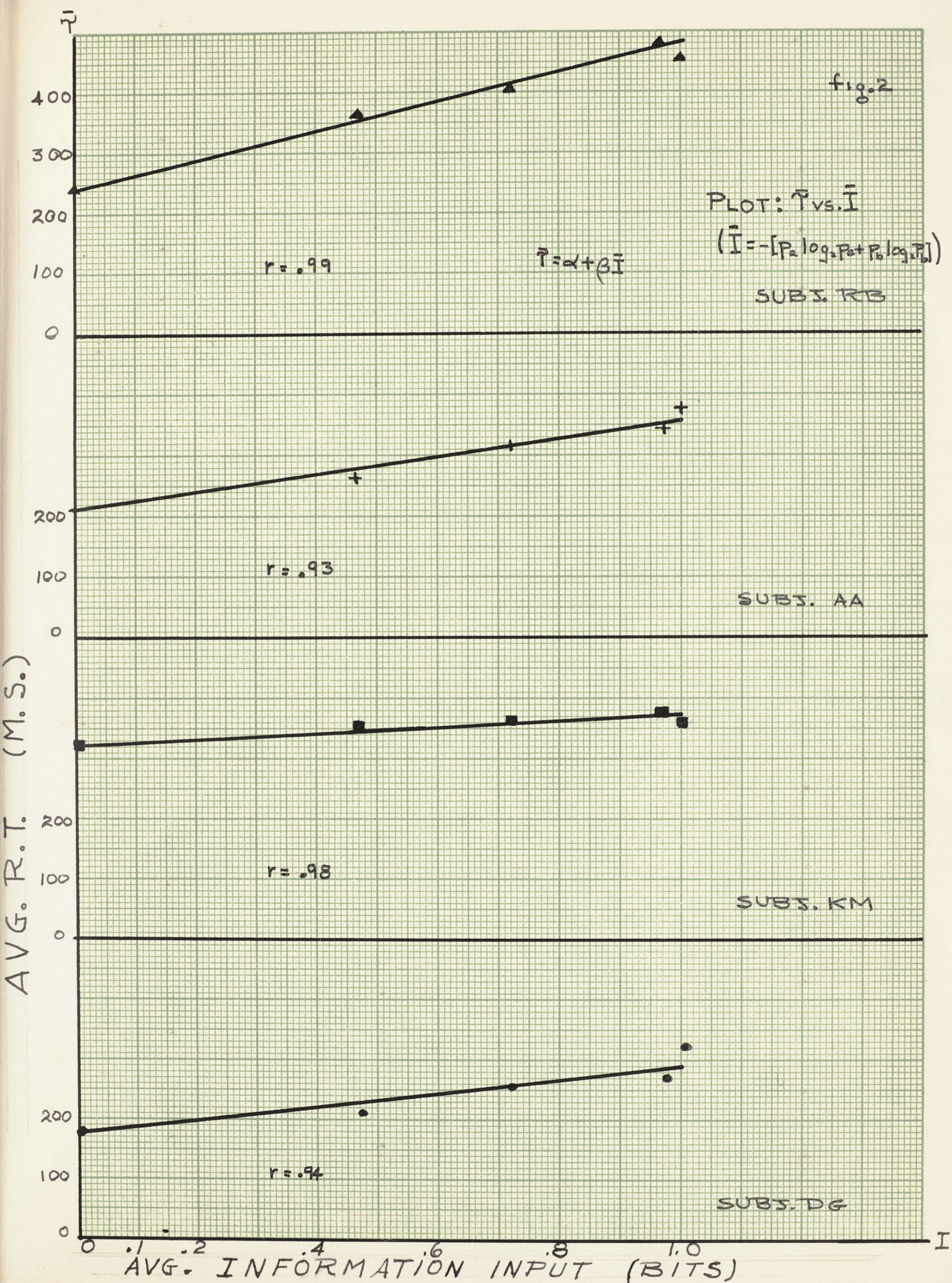


fig. 2
PLOT: $\bar{\tau}$ vs. $\bar{I}$
$(\bar{I} = -[P_a \log_2 P_a + P_b \log_2 P_b])$
$\bar{\tau} = \alpha + \beta \bar{I}$
r = .99
SUBJ. RB
r = .93
SUBJ. AA
r = .98
SUBJ. KM
r = .94
SUBJ. DG
AVG. R.T. (M.S.)
AVG. INFORMATION INPUT (BITS)

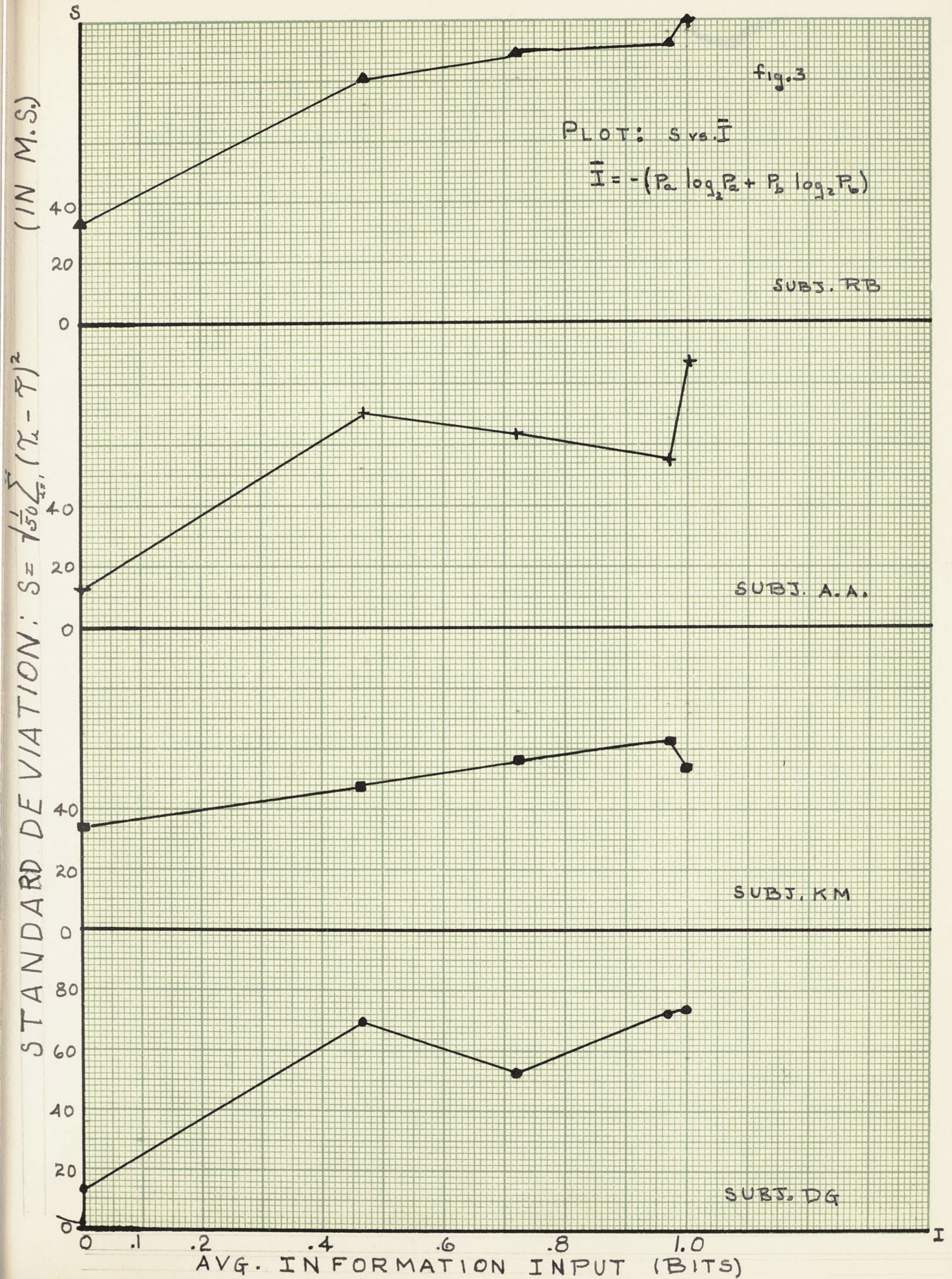
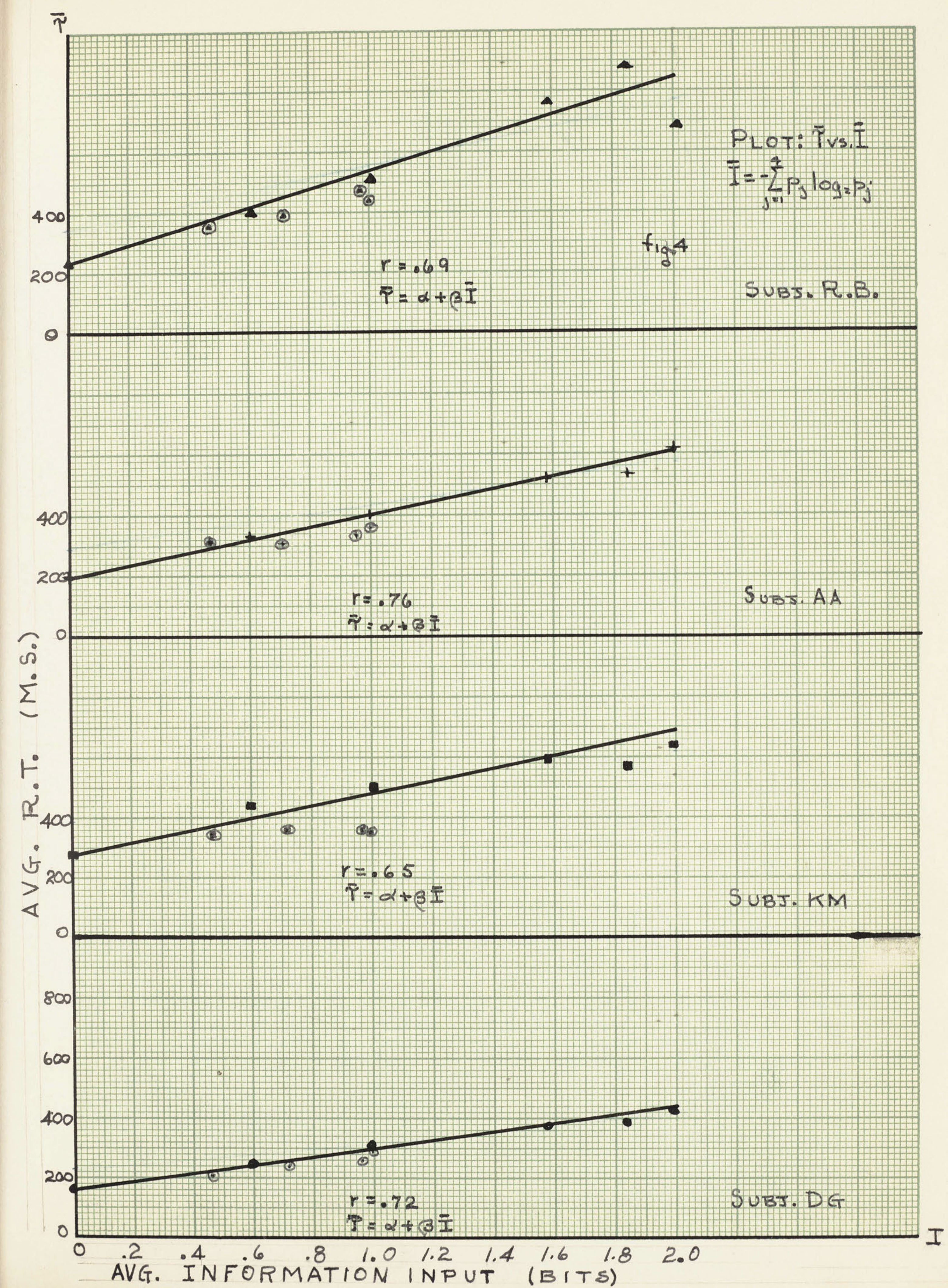
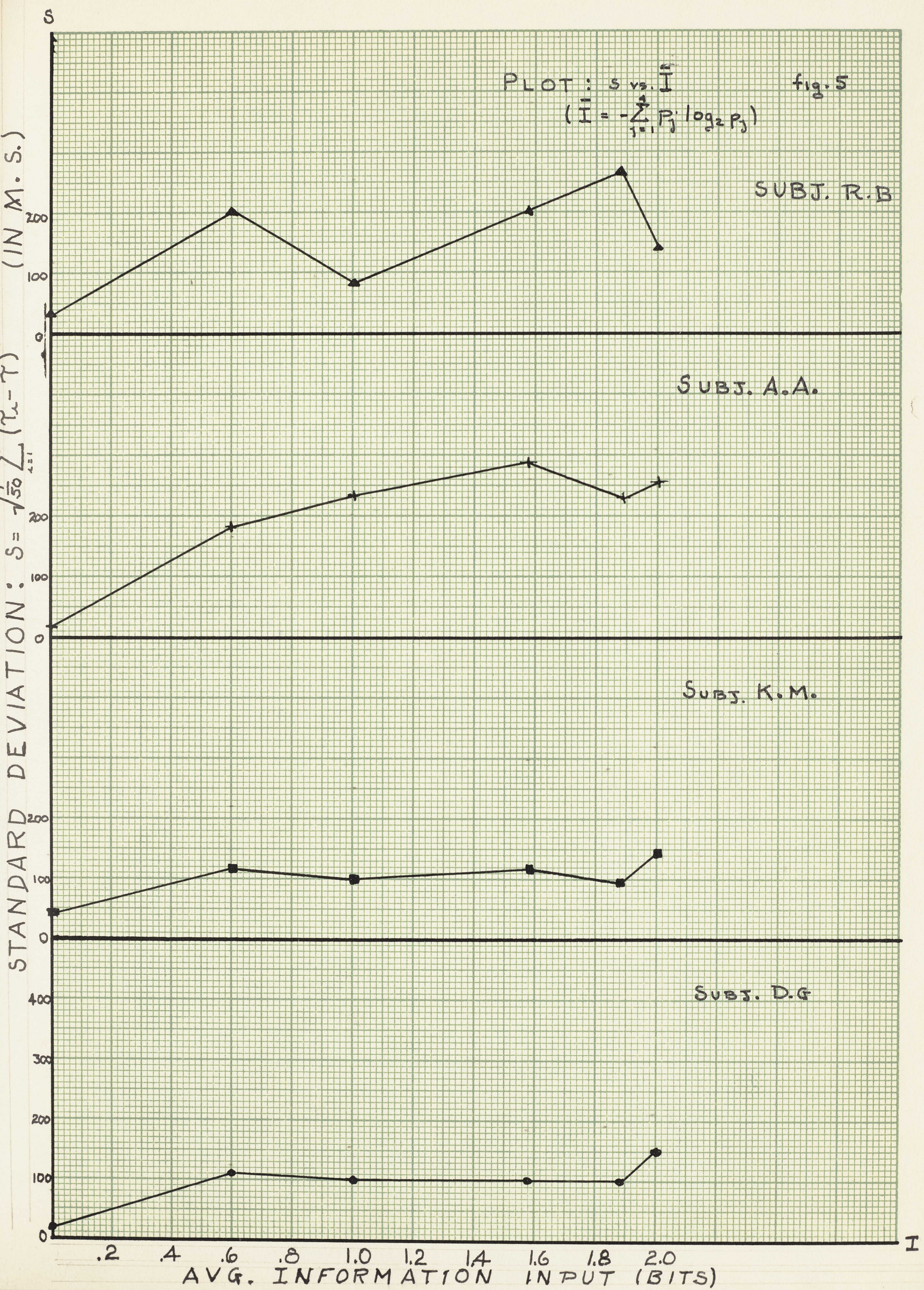


fig. 3
PLOT: S vs. $\bar{I}$
$\bar{I} = -(P_a \log_2 P_a + P_b \log_2 P_b)$
SUBJ. RB
SUBJ. A.A.
SUBJ. KM
SUBJ. DG
STANDARD DEVIATION: $S = \sqrt{\frac{1}{50}\sum_{i=1}^{50}(\tau_i - \bar{\tau})^2}$ (IN M.S.)
AVG. INFORMATION INPUT (BITS)
S
I
0 .1 .2 .4 .6 .8 1.0
0 20 40
0 20 40
0 20 40
0 20 40 60 80



τ̄
PLOT: τ̄ vs. Ī
Ī = -Σ pj log2 pj
fig. 4
r = .69
τ̄ = α + βĪ
SUBJ. R.B.
r = .76
τ̄ = α + βĪ
SUBJ. AA
r = .65
τ̄ = α + βĪ
SUBJ. KM
r = .72
τ̄ = α + βĪ
SUBJ. DG
AVG. R.T. (M.S.)
0
200
400
600
800
.2
.4
.6
.8
1.0
1.2
1.4
1.6
1.8
2.0
I
AVG. INFORMATION INPUT (BITS)



PLOT : S vs. $\bar{I}$
fig. 5
$(\bar{I} = -\sum_{j=1}^{4} p_j \log_2 p_j)$
SUBJ. R.B
SUBJ. A.A.
SUBJ. K.M.
SUBJ. D.G
STANDARD DEVIATION: $S = \sqrt{\frac{1}{50}\sum_{i=1}^{50}(\tau_i - \bar{\tau})^2}$ (IN M.S.)
S
I
0
100
200
300
400
.2
.4
.6
.8
1.0
1.2
1.4
1.6
1.8
2.0
AVG. INFORMATION INPUT (BITS)

Discussion

Since the mean RT's appear to fall along a line of regression $\bar{t} = \alpha + \beta \bar{I}$, it would be of value to be able to construct confidence limits on α & β. Before we attempt this, let us examine the conditions necessary to the success of such a task; then we shall examine our experiment to see if it falls within the realm of the method.

The conventional theory of confidence limits on lines of regression hinges on two assumptions:[2,3,4] First, that the dependent variable ($\bar{t}_i$) is independently, normally distributed about a true regression line $\bar{t}' = \alpha + \beta \bar{I}$. Secondly, that each $\bar{t}_i$ has the same variance (v^2). We shall examine the validity of these assumptions in view of the experiment at hand.

In a report published by the Controls Systems Laboratory of the University of Illinois, Christie and Luce[5] hypothesized that a reaction time (t) which entails a decision-making process, can be thought of as consisting of two components: $t = t_b + t_c$. t_b is a base time, dependent only upon the mode of stimulus presentation and motor actions required. Specifically, it is not dependent upon the character of the choice demanded. The value of t_c, in turn, depends only upon the choice demanded and not upon the stimulus mode or motor actions required. These assumptions seem, to the writer, to be reasonable ones, and the bulk of this discussion will depend upon them.

It is also reasonable to further restrict the nature of the base reaction time to be equivilent (in essence) to the simple reaction time corresponding to that mode of stimulus and that particular motor action demanded of the subject. For example, suppose the experimental situation is the two choice one, where either tone "a" or tone "b" can occur. Then t_b would be the simple reaction time associated with tone "a" if tone "a" were the stimulus and t_b would be the simple reaction time associated with tone "b" when tone "b" was the stimulus.

Classic experiments have shown that simple reaction times of this sort are normally distributed.[6] Hence, we make the following assertion: 1) t_b is a normal variate with mean M and variance V^2.

The distribution of t_c is a little less easy to come by. Christie and Luce[7] suggest the following point of view: "Suppose no decision has been reached by time t following stimulation at time 0. Then the probability that the decision will be reached between t and $t+\Delta t$ (where Δt is small) is approximately proportional to Δt, the constant of proportionality being λ."

It should be stressed that this is not a statement of fact, but merely a hypothesis which poseses the virtues of mathematical simplicity and intuitive reasonableness. Experiments have been done along the line of Luce & Christie's hypothesis, and the results are such that the hypothesis seems to be close enough to the right track not to warrant rejection.[8]

The preceding assumption, couched in probability terms leads us to the distribution of t_c (see appendix I):

$$f(t_c) = \lambda \exp(-\lambda t_c)$$

If the distributions of the respective variables t, t_b and t_c are denoted by h(t), $g(t_b)$ and $f(t_c)$ then

$$h(t_b + t_c) = h(t) = \int_0^t g(x)\, f(t-x)\,dx$$ [9]

This turns out to be a messy integral and it suffices to say that h(t) is not normal (see appendix II). Nor, for that matter is the distribution of $\bar{t}$, $p(\bar{t})$, normally distributed (see appendix III).

If the initial assumptions about the nature of decisien reaction times are accepted, we are led to the conclusion that the distribution of the mean reaction times is <u>not</u> normally distributed.

As was mentioned previously, the theory of confidence intervals for lines of regression hinges on two assumptions: That of normality and that of equality of variances. The first of these assumptions is violated if the preceding argument is accepted. However, <u>even if the first were not false</u>, the second would be violated by the results in tables 5 and 6. For if $\bar{t}_1, \bar{t}_2, \ldots, \bar{t}_k$ <u>were</u> normally distributed, then, for the given experimental conditions (equal number of samples) $\left(\frac{s_i}{s_j}\right)^2$ would be distributed* according to Snedcor's F distribution with 49 and 49 d.f. [10] If the hypothesis $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2$ is to be satisfied, then $\left[\frac{\max_i (s_i)}{\min_i (s_i)}\right]^2 < F_o \quad (i = 1,2..k)$ where $F_o = 1.61$ at the 5% point.

* for $i \neq j$, of course

Computations show that none of the data satisfy this criterian. Thus, construction of confidence intervals on the lines of regression, using conventional techniques, would be erroneous and the attempt is necessarily abandoned.

The reader may balk at the reasoning behind the first part of this argument. Does it make sense to accept a comparatively unproven hypothesis in preference to data at hand?

If the hypothesis is the only evidence to the contrary, certainly not. But, if there is, in addition, other empirical evidence to negate a potential assertion, then it is fair to summon all weapons (intuitive ones included) to make the contra-hypothesis more acceptable. (This is this writer's own view.)

In view of these developments, the writer has adopted a slightly different point of view in determining α & β than is ordinarily found in work of this sort. If the reader will examine figs. 2 & 4, he will notice that, the way the lines of regression are constructed, they always pass through the ordinate axis at the experimentally determined value of $\bar{t}$ corresponding to $\bar{I} = 0$. This was done intentionally, since, the value of $\bar{t}$ when $\bar{I}$ is 0 corresponds to a simple reaction time. These simple reaction times have considerably lower variances than have the decision RT's. Hence, if there are any "most reliable" points on figs. 2 & 4, these are certainly the ones. The slope β was found by the usual least squares method.

When the experiment was first conceived, it was thought that the study of errors would prove fruitful. If the subject makes an incorrect response, the information that he transmits is less than the information he receives. It seemed reasonable to assume that as the task becames more complex, the subject would make more mistakes. Stated in information theory language: The higher the information input, the higher the (probable) degree of equivocation.

This is a rather general statement, so let us qualify it in terms of the experiment at hand: The subject receives a stimulus with which is associated (on the average) $\bar{I}$ bits of information. He must make a decision, then react by giving the correct response. If there is a source of equivocation in this channel, one is justified in asking about its origins. As this writer sees it, the source is twofold: If the subject were not constrained to respond in as short a time as possible (i.e. if this were not an RT experiment) the source of equivocation would be strictly dependent upon the nature of the stimuli. If there were two stimuli, the subject would be less apt to make an error than if there were twenty.

However, this experiment demands that the subject "make up his mind" quickly. He is rushed, so to speak. Hence, time becomes a parameter influencing the information lost in the transmission process.

In our experiment, the possible number of stimuli never exceeded four. These four stimuli were carefully chosen so that, given a chance to think about them, a person with normal hearing would not confuse them. Any two tones in the stimulus ensemble were at least an octave apart; octave confusion was minimized, in turn, by choosing the tones at 200, 750, 2500 and 6000 cps. It was hoped, that once this source of error was minimized, equivocation would be almost exclusively dependent upon RT. As it turned out however, the trend for the subjects as a whole was to make only a small number of mistakes so that $\bar{I}_{in} \approx \bar{I}_{transmitted}$. (See tables on pages 10 and 11.) Hick[11], in an RT experiment, urged his subjects to react faster, even if it cost them some accuracy. It did, and he found that $\bar{t}$ was reasonable correlated with $\bar{I}_{transmitted}$.

Summary

Two difficulties were encountered in analyzing the data. The first concerned the nature of the distribution of $\bar{t}$. This distribution (once a few assumptions were made) turned out to be non-Gaussian, and rather cumbersome, to boot. As a result, any statistical treatment of the data was sorely restricted. One might be tempted to use the Central Limit Theorem and thus arrive at a normal distribution for $\bar{t}^{-1/2}$. However, in this case, the use of large sample methods on a sample of size n=50 is a risky business. Moreover, the primary use of normality would be the construction of confidence limits about the parameters in the linear regression scheme. To accomplish such a task using conventional methods, a condition of homoscedasticity (equality of variances) must be realized. It was not. These two factors combined to make the usual confidence interval arguments lose their utility.

The second disappointment was encountered when, contrary to "pre-experimental intuition", errors did not display a meaningful trend.

The experiment is not without meaning, however. Bricker, in his contribution to Quaster's survey on information theory methods in psychology, states that modern work in the region of RT experiments has suggested the hypothesis [13] "...that average RT, unless otherwise restricted, is an increasing linear function of the average amount of information transmitted from the display to the response." (In our case, the average information transmitted was essentially equal to the average information received by the subject.) Furthermore, Bricker poses the question: Does this relation hold no matter how $\bar{I}$ is varied? Hyman[14], using visual stimuli, showed this to be so over a range of information values from 0 to 3 bits.

By consulting fig.4, we see that the points which are encircled, fall, in general, quite close to the line of regression of $\bar{t}$ on $\bar{I}$ for the four stimulus case. These encircled points are the mean RT's corresponding to the average information inputs for the two choice case. Thus, (over a relatively small range of $\bar{I}$ values), our data seems to substantiate this hypothesis.

In view of the intial objectives of this investigation- to investigate the dependence of decision reaction times upon the average information associated with the stimulus to which S is responding- some degree of suceess has been achieved. However, there is room for improvement within the basic structure of this experiment.

The variabilities for all points show fluctuations which might well be reduced if the training periods for each section of the experiment were extended considerably. In addition, the number of points which constitute a set, could well be extended beyond 50. This would both increase the reliability of the experiment and the validity of using the Central Limit Theorem in dealing with the variable $\bar{t}$.

It would also be of interest to extend the range of information inputs to five, six or seven bits, if possible. It will be a difficult task, however, using auditory displays; to handle five bits of information requires the ability to discriminate between 32 alternatives. In auditory displays, this is no mean accomplishment.

Finally, the question of variability should come under scrutiny. In our experiment, for a fixed value of $\bar{I}$, the variability was higher in the four alternative case than in the two alternative case. Was this caused by subject variability, or is the variance really dependent upon the number of alternatives (which may not be unique for a fixed $\bar{I}$)?

The answer to these questions (and more important, the answer to the question: Are they important questions?) lies in future experimentation.

Conclusions

If the information input is varied in a situation where there are two possible stimuli, the line of regression of $\bar{t}$ on $\bar{I}$ displays a great deal of linearity with correlation coefficients between $\bar{t}$ and $\bar{I}$ quite close to unity.

A similar, but weaker statement can be made concerning the four choice situation.

When the two lines of regression are superimposed upon each other, they appear to coincide for three of the four subject. To some degree, this substantiates a hypothesis that average RT's are increasing linear functions of average Information transmitted, no matter how the value of $\bar{I}$ is determined (e.g. independent of the number of alternatives).

Appendix I

Let $F(t_c) = \int_0^{t_c} f(\tau)\,d\tau$;

$$1 - F(t_c) = \int_{t_c}^{\infty} f(\tau)\,d\tau$$

$f(\tau)\,d\tau = Pr\{\tau \le t_c \le \tau + d\tau\} = dF(\tau)$

Let $\{AB\}$ be the occurrence of an event in $(\tau, \tau + d\tau)$ after time τ has elapsed.

Let $\{B\}$ be the occurrence of an event after time τ has elapsed.

Let $\{A|B\}$ be the occurrence of an event in $(\tau, \tau + d\tau)$ given that time τ has elapsed without the occurrence of the event.

Then

$$Pr\{A|B\} = \frac{Pr\{AB\}}{Pr\{B\}} = \lambda\,d\tau$$

$$Pr\{AB\} = dF(\tau)$$

$$Pr\{B\} = 1 - F(\tau)$$

$$\therefore \frac{dF(\tau)}{1 - F(\tau)} = \lambda\,d\tau$$

$$-\log_e c(1-F) = \lambda\tau$$

$$c(1-F) = e^{-\lambda\tau}$$

$$F(\tau) = 1 - \frac{1}{c}e^{-\lambda\tau}$$

$$F(0) = 0 \Longrightarrow c = 1$$

$$F(\tau) = 1 - e^{-\lambda\tau}$$

$$dF(\tau) = f(\tau) = \lambda e^{-\lambda\tau}$$

Appendix II

Suppose two independent variables x + y have distributions with MGF's $M_y(\theta)$, $M_x(\theta)$.

Let x be normally distributed with mean a_x, variance v_x^2.

Then $M_x(\theta) = e^{a_x \theta} e^{-1/2\, \theta^2 v_x^2}$

$$M_{x+y}(\theta) = M_x(\theta)\, M_y(\theta)^{14}$$

Suppose the distribution of $z = x + y$ has MGF $M_z(\theta)$.

<u>Assert</u>: $f(z)$ is normal $\Leftrightarrow$ $g(y)$ is normal.

$\Rightarrow$:

$$M_z(\theta) = e^{a_z \theta} e^{-1/2\, \theta^2 v_z^2}$$

$$M_y(\theta) = \frac{M_z(\theta)}{M_x(\theta)} = e^{(a_z - a_x)\theta} e^{-1/2\, \theta^2 (v_z^2 - v_x^2)}$$

$$\doteq e^{a_y \theta} e^{-1/2\, \theta^2 v_y^2}$$

$\Leftarrow$:

$$M_y(\theta) = e^{a_y \theta} e^{-1/2\, \theta^2 v_y^2}$$

$$M_z(\theta) = M_x(\theta) M_y(\theta) = e^{(a_x + a_y)\theta} e^{-1/2\, \theta^2 (v_z^2 + v_x^2)}$$

$$= e^{a_z \theta} e^{-1/2\, \theta^2 v_z^2}$$

Since there is a one-one correspondence between MGF's and PDF's[15], (for M analytic in a region containing the origin) the assertion is proved.

Appendix III

$P(\bar{t}) = P(\bar{t}_b + \bar{t}_c)$

$\bar{t}_b$ is normal with mean m and variance $\frac{v^2}{m}$

$$M_{\bar{t}_c}(\theta) = M_{\frac{t_{c_1} + \cdots + t_{c_m}}{m}}(\theta) = M^m_{t_c}(\theta/m)$$

$$M_{t_c}(\theta) = \int_0^\infty e^{\theta\tau} \lambda e^{-\lambda\tau} d\tau = \frac{\lambda}{\lambda - \theta}$$

$$M^m_{t_c}\left(\frac{\theta}{m}\right) = \left(1 - \frac{\theta}{\lambda m}\right)^{-m} = M_{\bar{t}_c}(\theta)$$

This is the MGF of a Pearson type III distribution.

$\therefore$ By previous argument $P(\bar{t})$ is not normal.

5,7,8: Christie, L.S. & Luce, R.D. "Suggestions for the Analysis of Reaction Times and Simple Choice Behavior." University of Illinois Control Systems Laboratory Report # R-53 April 1954. (5-pg.9; 7-pg.15; 8-pg.3)

4: Cramér, H. Mathematical Methods of Statistics. Princeton University Press. 1954 (pg.548)

11: Hick, W.E. "On the Rate of gain of Information." Quart. Journal of American Psychology. 1952. (pp.11-26)

3,10,12 : Hoel, P.G. Introduction to Mathematical Statistics Second Edition. John Wiley & Sons Inc. (3-pg.229; 10-pg.235; 12-pg.107)

1,14: Hyman, R. "Stimulus Information as a Determinant Of Reaction Times." Journal of Experimental Psychology. vol. 45, no. 3, 1953 (pp.188-196)

13: Quastler, H. Information Theory in Psychology; Problems and Methods. Free Press Publishers, Urbana, Ill. 1955 (pp.350-359)

2: Wilks, S.S. Mathematical Statistics. Princeton University Press.1950 (p.170)

9: Woodward, P.M. Probability and Information Theory, With Applications to Radar. McGraw Hill Book Co. N.Y. 1955.(p.13)

6: **Woodworth, R.S. and Schlosberg, H. Experimental Psychology. Revised Edition. Henry Holt and Co., New York. 1954.**