C.P. No. 1030

C.P. No. 1030

MINISTRY OF TECHNOLOGY

AERONAUTICAL RESEARCH COUNCIL

CURRENT PAPERS

The Effects of Extended Practice on Performance in a Tracking Task

By

R. C. Hornby, B.Sc. and R. Wilson, B.Sc., Ph.D.
Deportment of Aeronautical Engineering,
The Queen's University of Belfast

LONDON: HER MAJESTY'S STATIONERY OFFICE

1969

Price 4s 6d. net

C.P. No. 1030*

December, 1967

The Effects of Extended Practice on Performance in a Tracking Task

-By -

R.C. Hornby, B.Sc. and R. Wilson, B.Sc., Ph.D.
Department of Aeronautical Engineering,
The Queen's University of Belfast.

SUMMARY

Experimental measurements of human controller performance have been made during extended periods of practice in visual sine-wave tracking tasks. It has been found that, irrespective of task difficult, RMS error scores decreased to such small magnitudes that differences in scores due to different task variables would have no practical significance. Thus the averaged values of steady scores when tasks are well learnt are meaningless for subject or task difficulty comparisons. It has been shown that the performance scores vary in an exponential manner with the number of task repetitions and it is proposed that an empirical constant related to the rate of decrease of scores be used as a measure of relative task difficulty.

Introduction

Performance scoring is a method of measuring the proficiency of a human pilot in a manual control task. Often this kind of experimental approach is used to compare the relative skills of different controllers, or the relative merits of different controls systems. Experience has led to the notion that there exist "plateaus of performance" for both subject variations and task variations. Thus, if different subjects were to practice a particular control task for a period of time, it is assumed that the performance scores would eventually reach steady values as shown in Fig. 1.

During the first N_1 repetitions the scores of both subjects are continually decreasing. During the repetitions numbered N_1 to N_2 the scores of the subjects A and B are the constant values S_A and SB. During the repetitions numbered N_2 to N_3 the respective scores are unchanged.

In order to draw significant conclusions from such a set of data, two statistical tests are required. The first test is required to show that the mean values of the scores in the interval N2 to N do not differ significantly from the mean values in the interval N_1 to N_2. (In practice, measured human performance scores would not decrease in monotonic order as shown in Fig. 1 because of the random variability of human controller behaviour. Neither would the scores in the steady interval N_1 to N_3 be exactly constant, hence the need for statistical tests.)

A/

---

*Replaces A.R.C.29 887

A second test is required to determine if there is any significant difference between S_A and S_B when the first test shows that no further decrease of averaged score is taking place. If the difference between S_A and S_B were found to be significant then it would be concluded that subject B was a more skilful controller than subject A.

In a similar manner, two different control tasks could. be compared on the basis of the performance scores of a particular subject or group of subjects. If a significant difference between the mean steady scores in the two tasks were found it would be concluded that one task was more difficult than the other.

This procedure has generally proved useful as a rough and ready guide for the selection of personnel for machine operation and for the design of dynamic systems. Nevertheless it has shortcomings and there is still a need for improved methods which take account of the effects and potential effects of operator learning on control performance. An increasing amount of doubt has been expressed as to whether plateau levels of performance actually occur, or whether scores would continue to decrease on the average if lengthy practice were continued.

Moreover, there is a class of problems in which the performance scores, e.g., error scores in a tracking task, can become so small that they are difficult to measure accurately. Moreover, provided the measuring apparatus is designed to resolve the smallest difference which has any practical relevance, nothing is to be gained by introducing more sopnisticated and more accurate measurements. Any difference detected in the mean scores which might be significant in the statistical sense could be trivial in relation to the task at hand.

The experiments reported in the present paper have illustrated these kinds of phenomena in sine wave tracking tasks. Extended periods of practice have shown that apparently steady plateau levels of performance are of relatively short-term duration end that the error scores continued to decrease on the average as practice progressed. Eventually the scores reached the limitations of the measuring apparatus. Had a more accurate measuring procedure been adopted then any differences which might have been detected among the independent variables would have been trivial.

Thus the state of affairs was that two subjects performed six tasks of different degrees of difficulty, which included twelve independent variables. The final steady mean performance scores were the same in all cases after sufficient periods of practice. It is clear that comparisons among the independent variables cannot be based on final steady performance scores. It is shown, nevertheless, that a significant comparison can be based on the average rate of change of scores with the number of task repetitions.

Experimental Description

The subjects sat in a small enclosed cockpit watching the motion Of a spot of light on a 5 in. diameter oscilloscope at a distance of about 24 in. Control was effected with a 20 in joy-stick having low friction and linear spring force gradients of about 1 lb/in, An electrical output from the stick was fed to an analogue computer and control dynamics relating spot movement to stick movement were:

(i) position control, $x = 2.0$

(ii) velocity control, $\frac{dx}{dt} = 4.0$

(iii) acceleration control, $\frac{d^2x}{dt^2} = 4.0$

where/

where x = spot displacement in cm from a reference point

= stick deflection in inches, measured from centre at the hand grip.

Performance scores were measured in both single axis and two axes tasks. In the single axis tasks a sinusoidal input of 0.2 cps nominal frequency was fed to the horizontal plates of the oscilloscope. If this motion were uncontrolled the maximum amplitues of the spot from the reference would have been ±2.5 cm. The subjects attempted to track the signal by controlling the spot to keep it as close as possible to the reference,

In the two axes tasks an additional sine wave input was applied to the vertical plates of the oscilloscope. This input had the same amplitude and frequency as the horizontal input but was displaced in phase by 90°. The uncontrolled motion of the spot was thus a circular path of 2.5 cm radius.

Individual tracking thus occupied one minute. During the first 15 sec. no scoring occurred. In the next 30 sec. the tracking error voltage was squared and integrated in the computer, There was then a rest period of 15 secs. Two subjects alternately performed sessions of ten one-minute runs. At the end of a run the final voltage corresponding to the total integrated squared error was recorded on a trace recorder. The traces were later read to the nearest whole number percentage of RMS error/RMS input. A 100% score would, therefore, be obtained with an RMS error of 1.8 cm on the oscilloscope, a 10% score with 0.18 cm and 5% with 0.09 cm, etc,

The main object of the tests was to observe the effects of extended practice on human operator performance in tracking tasks of known difficulty. A sine wave input was chosen for two reasons. First, its simple repetitive form could be learned quickly by the subjects, thus helping to reduce the large number of laborious task repetitions which would be required in any case. Secondly, by using this same input throughout it could be assumed that the difficulty of the tasks depended only on the nature of the control dynamics and the number of axes being controlled. Thus we could, with reasonable certainty, estimate the relative difficulties of the tasks in qualitative terms which could then be compared with the quantitative results. It is known that velocity tracking is a little more difficult than position tracking, and acceleration tracking considerably more difficult than either, Corresponding two axes tasks are more then twice as difficult as single axis tasks, the controller has to exert twice the effort to control about both axes simultaneously, and as well as this he has to make additional decisions as to which control movements are required for which axis.

In all cases, task repetitions were continued until scores of 5% were achieved regularly (i.e., at least five times in a session of ten repetitions). A 5% score implied a root-mean-squared error during the run of less than a millimetre, and it was apparent that errors as small as this could easily have been introduced by parallax in the display. Hence, as pointed out in the introduction to this report, any differences in mean steady scores of this order of magnitude would be classed as "trivial" in relation to the nature of the task. Moreover, it would not have been possible to measure scores lower than this value with any great accuracy. For a displayed error of 1 mm the voltage output of the servo-multipliers which were used for squaring was about 1/10th of a volt. After integrating small voltages of this order for half a minute the overall voltage is also contaminated by effects such as amplifier drift and d.c. bias, Thus it seemed logical to disregard scores apparently lower then 5%. These were read as 5%.

Experimental/

Experimental Results

Variations in performance scores with the number of task repetitions are shown in Pigs. 2 to 8. Figs. 2 end 3 are the results obtained with single axis position and velocity controls. The crosses represent the scores in each of the individual runs. Scores were reduced to the 5% level after about 50 to 70 repetitions.

The scores obtained with single axis acceleration dynamics are shown in Fig. 4. The crosses in this figure are the mean scores of every 10 task repetitions. With acceleration controls the subjects required over 500 task repetitions to improve performance to the 5% level.

In Figs. 5, 6, 7 and 8, which illustrate the variations of scores in the two axes situations, the crosses are the mean value of 10 vertical scores and the circles the mean values of 10 horizontal scors. The number of task repetitions required to reach the 5% level was more than double those required for corresponding single axis situations.

The fitting of learning curves

The plots of performance scores as functions of the number of task repetitions can be fitted with exponential curves of the form:

$$\bar{S} = Ae^{-\frac{n}{B}} + C \qquad \text{....(I)}$$

where $\bar{S} = \dfrac{\text{RMS error}}{\text{RMS input}} \times 10\%$

n = number of task repetitions

A, B, C = empirical constants.

Such curves are shown in Figs. 2 to 8 and the values of the appropriate constants are listed in Table I, These curves were obtained by plotting log $(\bar{S}-C)$ against N *and drawing a straight* line through the points by eye. The constants A and B for the curve shown were obtained from the equation of the line. It would, of course, have been possible to obtain the best fitting line more rigorously by submitting the experimental points to a computer programme but this refinement. was not considered necessary at this stage of the work. It may be adopted later. Due to the process of rounding off small scores to 5%, asymptotic convergence is ensured and the constant C in all oases is equal to 5. Although it may happen that ten or twenty repetitions will produce a score of 5% there is, nevertheless, still a finite probability that a score greater than 5% could be obtained in subsequent repetitions. Hence the convergence of the mean of the scores is, in fact, asymptotic and not absolute.

The constant A is a measure of the initial proficiency of the subject without practice, In effect the value:

$$\bar{S}_o = A + C$$

is the expected value of the subject's score for $n = o$. In a lengthy series of experimental tests there will inevitably be "transfer of training" effects, If a subject learns a task and the task is made slightly different by a change in

one/

one of the independent variables the subject may initially perform a good deal better in the new task than he would have done had he not had the previous training. In the present tests both subjects initially performed better in two axes tasks than they did in single axis situations even though two axes tasks are considerably more difficult than single axis tasks, This resulted from improvements in the subjects' overall skill due to practice in the single axis tasks which were performed first, It is felt that the value of A has little relevance to the degree of task difficulty.

Consideration of the values of the constant B averaged for the two subjects (Table II) shows that this number at any rate places the six tasks in the order of difficulty which the two operators were agreed upon.

Moreover this number is evidently insensitive both to the order of testing end to the subjects controlling ability. It is, therefore, proposed to adopt the constant B as a measure of task difficulty. If this is done, and taking the value of B in single axis position control as unity, the relative difficulties of the six tasks are as shown in Table III.

Idealised learning

In a difficult tracking task a subject's error scores decrease very slowly with practice. Thus ono of the difficulties in fitting learning curves is the fact that it might require an impractically large number of task repetitions to establish the trend of the curve. For example, the 1,500 or so task repetitions which were made by the subjects in the two axes acceleration control sine tracking task would be considered to be too many for a routine experiment in which many independent variables were to be tested.

Considerable saving of effort could be achieved if it were possible to predict a subject's ultimate performance from his early performance. If me assume that error scores in a tracking task will vary with task repetitions according to the law given by equation (I), then, by taking logarithms we see that:

$$\log(\bar{S}-C) = \log A - \frac{1}{B}\,n \qquad \ldots\ldots(2)$$

Thus if $\log(\bar{S}-C)$ is plotted against n we obtain a straight line of slope $-\frac{1}{B}$, where B is the index of task difficulty sought for. Hence it is only necessary to perform a few task repetitions to determine this slope with a reasonable degree of accuracy. In Pig. 9 the two axes acceleration scores of subject 1 have been plotted in the logarithmic form. The data points from the first few hundred task repetitions lie reasonably on a straight line. However, when this line is extrapolated we find that later scores do not lie on the line. The constant B obtained from the slope of the line in Fig. 9 has the value 263, as compared with the value of 400 obtained by fitting an exponential curve to the whole 1,500 pairs of scores.

The initial and more rapid rate of learning will be referred to as "idealised learning". When a human operator first attempts to perform a difficult task it is reasonable to assume that he concentrates his whole attention on the task, thus using up all his spare mental capacity. He achieves a certain score, S, but he has been controlling very inefficiently by making a lot of imprecise control movements. After a short time he makes fewer mistakes and can achieve the performance score S_1 with less than his maximum effort, which leaves him with an amount of spare mental capacity. He then exerts maximum effort again and the spare mental capacity enables him to decrease the error slightly.

Idealised/

Idealised learning is thought to be learning which occurs when the controller is continually using up all his spare mental capacity and always exerting maximum effort as his controlling ability improves. However, a controller learns at the idealised rate only during the initial practice period. When he had reduced the tracking errors to a level considerably smaller than his unpracticed level he no longer exerts maximum controlling effort. He prefers to continue controlling with increasing amounts of spare mental capacity rather than trying to improve his performance at a greater rate by exerting maximum effort. Actual learning curves will therefore depart further and further from idelised learning curves as practice progresses.

Nonetheless, the experimental results suggest that values of B for idealised learning curves may be in linear proportion to values of B for actual learning curves. Hence idealised learning curves may prove to be equally valid for comparative purposes. Table IV shows that the ratio $\frac{B_{2\ axes}}{B_{1\ axis}}$ for the acceleration control scores of subject 1 is approximately the same for both idealised learning and actual learning.

Conclusions

The following conclusions apply to the tracking of low frequency ($\sim$ 0.2 ops) sine waves. Whether they are valid for the tracking of more random input signals should be determine& by further tests.

(1) The RMS values of tracking errors computed during short controlling runs continued to decrease on the average as practice was extended.

(2) In six tasks of varying difficulty, ranging from single axis position tracking to two axes acceleration tracking, the only foreseeable convergent limit to the scores was that imposed by the display threshold.

(3) It is not practical to make comparisons among the different tasks on the basis of the magnitude of the scores when all improvement has ceased.

(4) Scores vary with task repetitions according to an exponential law of the form

$$\bar{S} = AE^{-\frac{N}{B}} + C$$

where N = the number of task repetitions

S = expected value of score

A, B = empirical constants

C = lowest score of practical significance (determined by display threshold or limitations of measuring apparatus).

(5) The constant A depends upon transfer of training effects and the subjects' initial controlling ability. It does not give any valid information regarding comparative task difficulty,

(6) The constant B is independent of transfer of training effects and the subjects' initial controlling ability and is a measure of relative task difficulty.

(7) The concept of idealised learning as described in the present report leads to a rapid determination of relative task difficulty which does not require a large number of laborious task repetitions.

TABLE I/

CONTROL DYNAMICS	SUBJECT 1			SUBJECT 2		
	A	B	C	A	B	C
I-AXIS POSITION	15	11	5 5	20	20	5
" VELOCITY				25		5
" ACCELN	40	150	5	40	126	5
2-AXES POSITION		40			33	5
" VELOCITY	10	83	5	16	56	5
" ACCELN	22	400	5	22	500	5

TABLE I

CONTROL DYNAMICS	B average
I-AXIS POSITION	15.5
" VELOCITY	22.5
" ACCELN	150
2-AXES POSITION	37.5
" VELOCITY	69.5
" ACCELN	450

TABLE II

CONTROL DYNAMICS	RELATIVETASK DIFFICULTY
1-AXIS FOSITION	1 .00
" VELOCITY	1.45
" ACCELN	9.68
2-AXES POSITION	2.42
" VELOCITY	4.49
" ACCELN	29.00

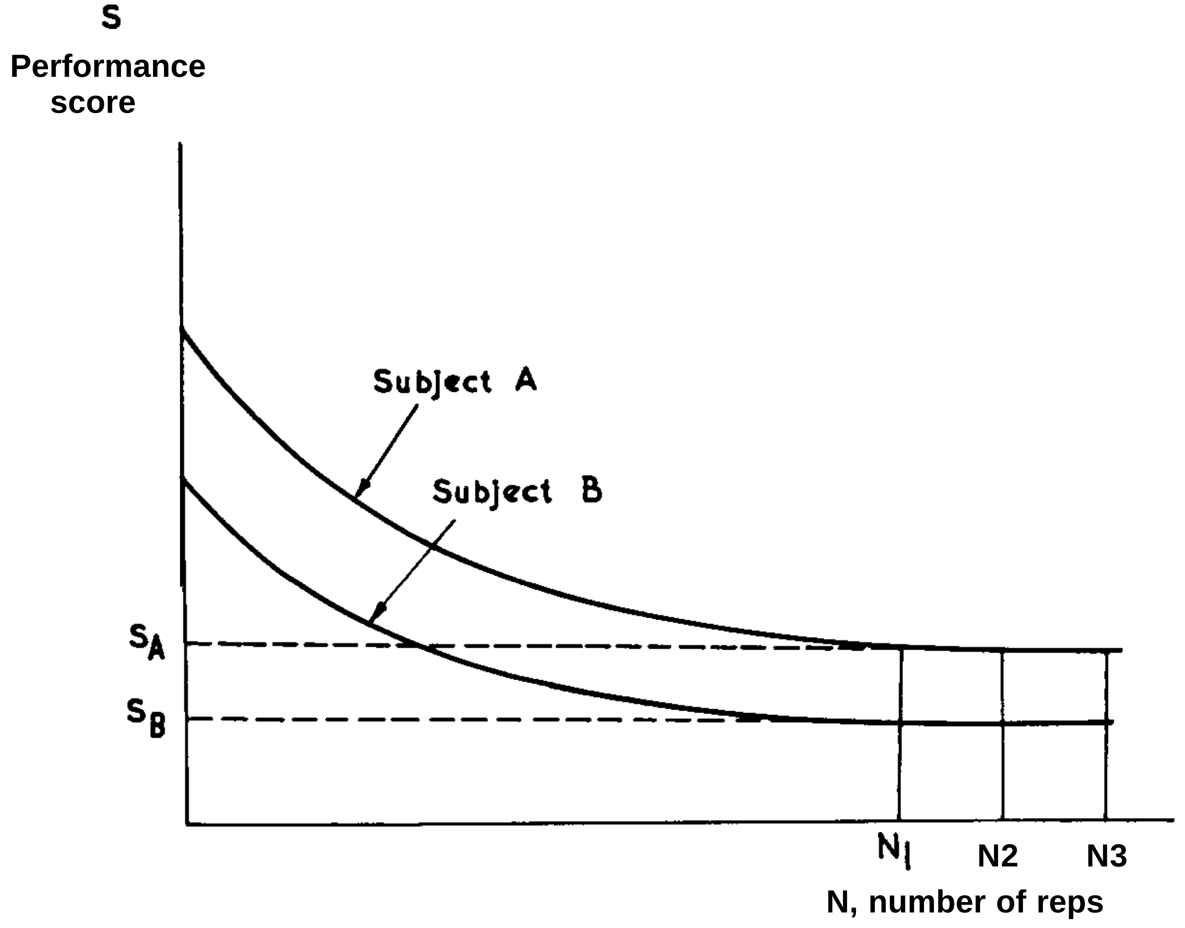
TABLE III

	1-AXIS ACCELN CONTROL. VALUE OF B FOR SUBJECT 1	24'S ACCELN CONTROL. VALUE OF B FOR SUBJECT 1	$\frac{B_{2\text{-axes}}}{B_{1\text{-axis}}}$
ACTUAL LEARNING	150	400	2.7
IDEALISED LEARNING	106	263	2.5

TABLE IV

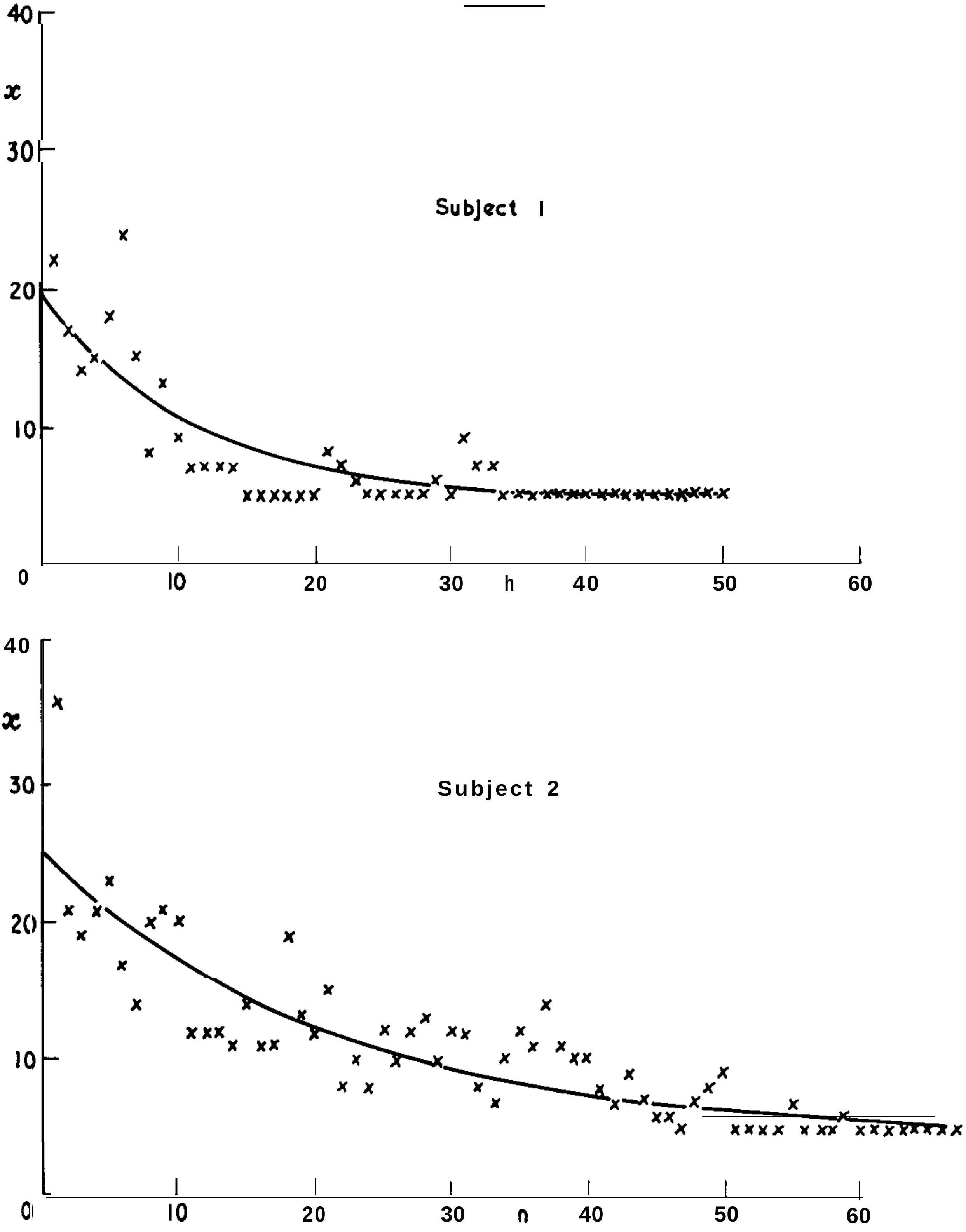
JED

FIG I



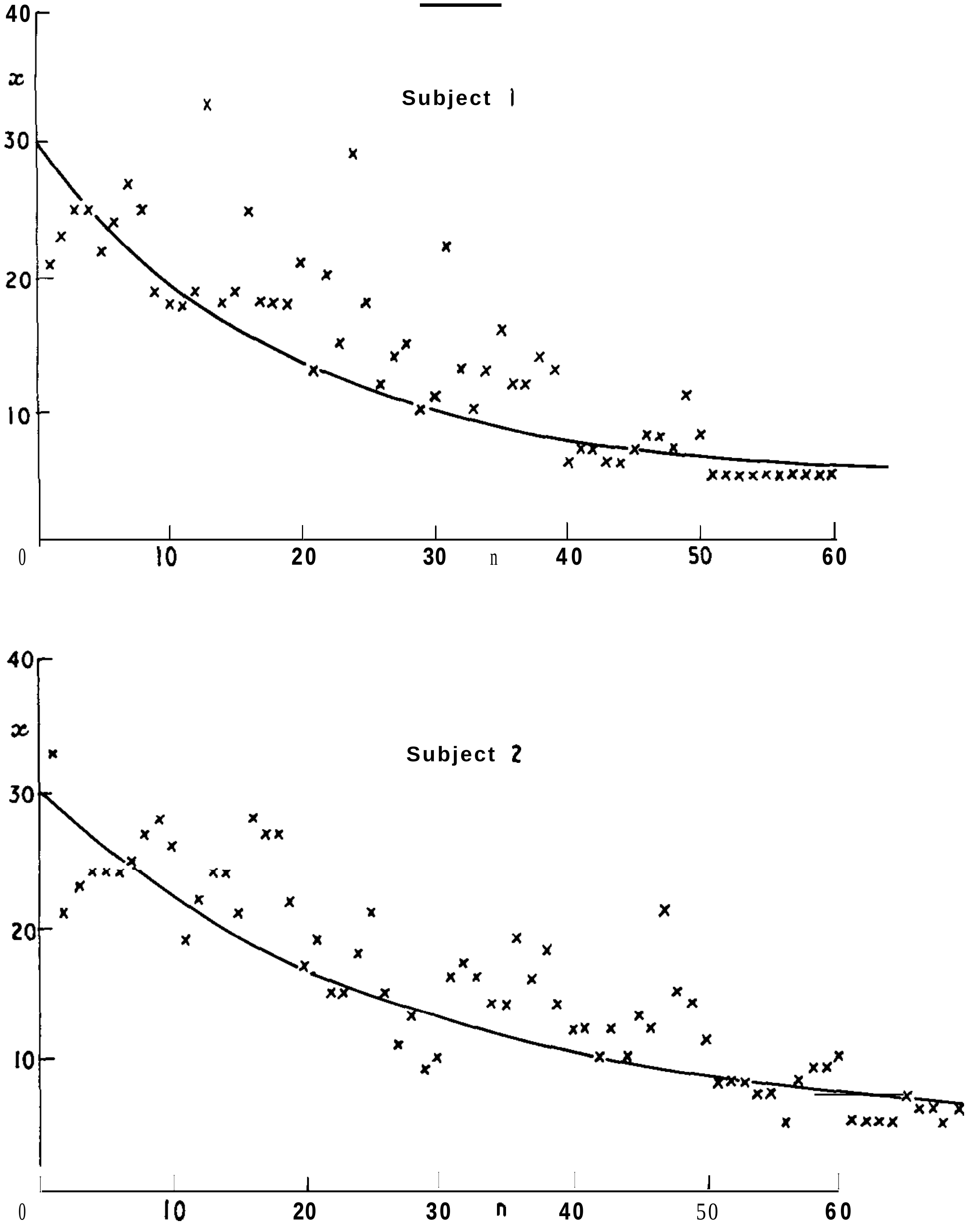
Hypothetical variations of performance scores with task repetition for two subjects, A and B

FIG. 2



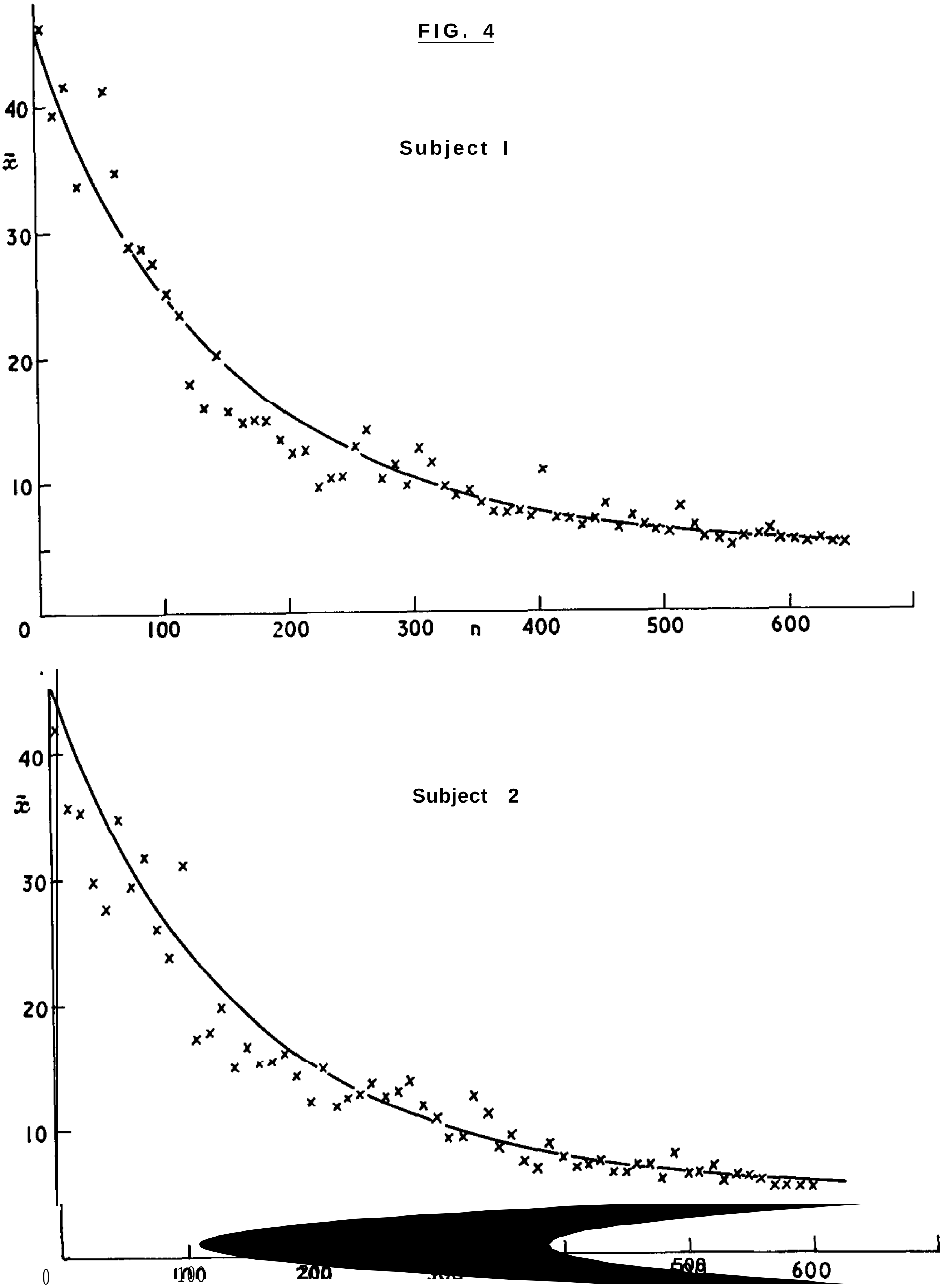
Learning curves. One-axis position control

FIG 3



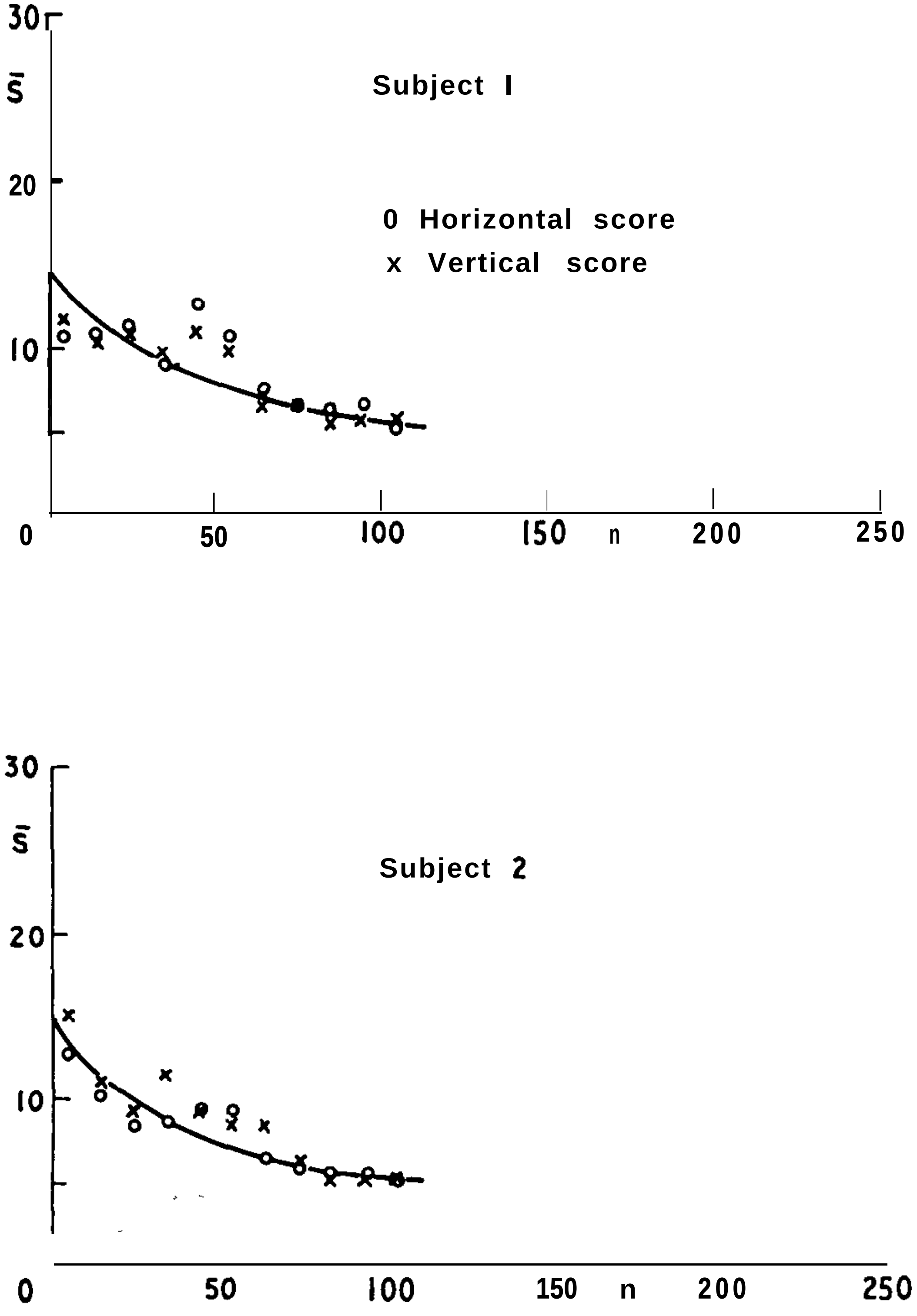
Learning curves One-axis velocity control

FIG. 4



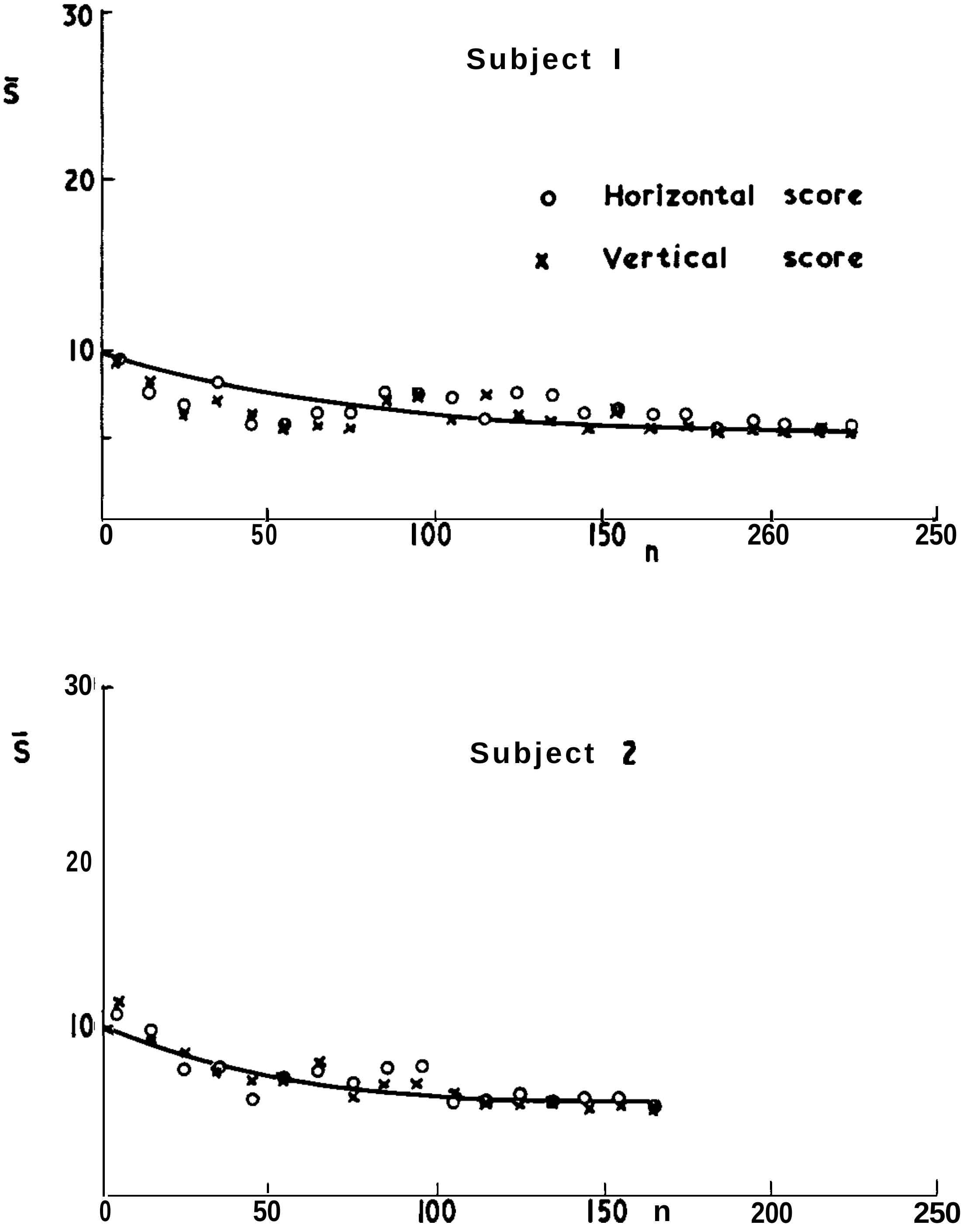
Learning curves. One-axis acceleration control

FIG. 5



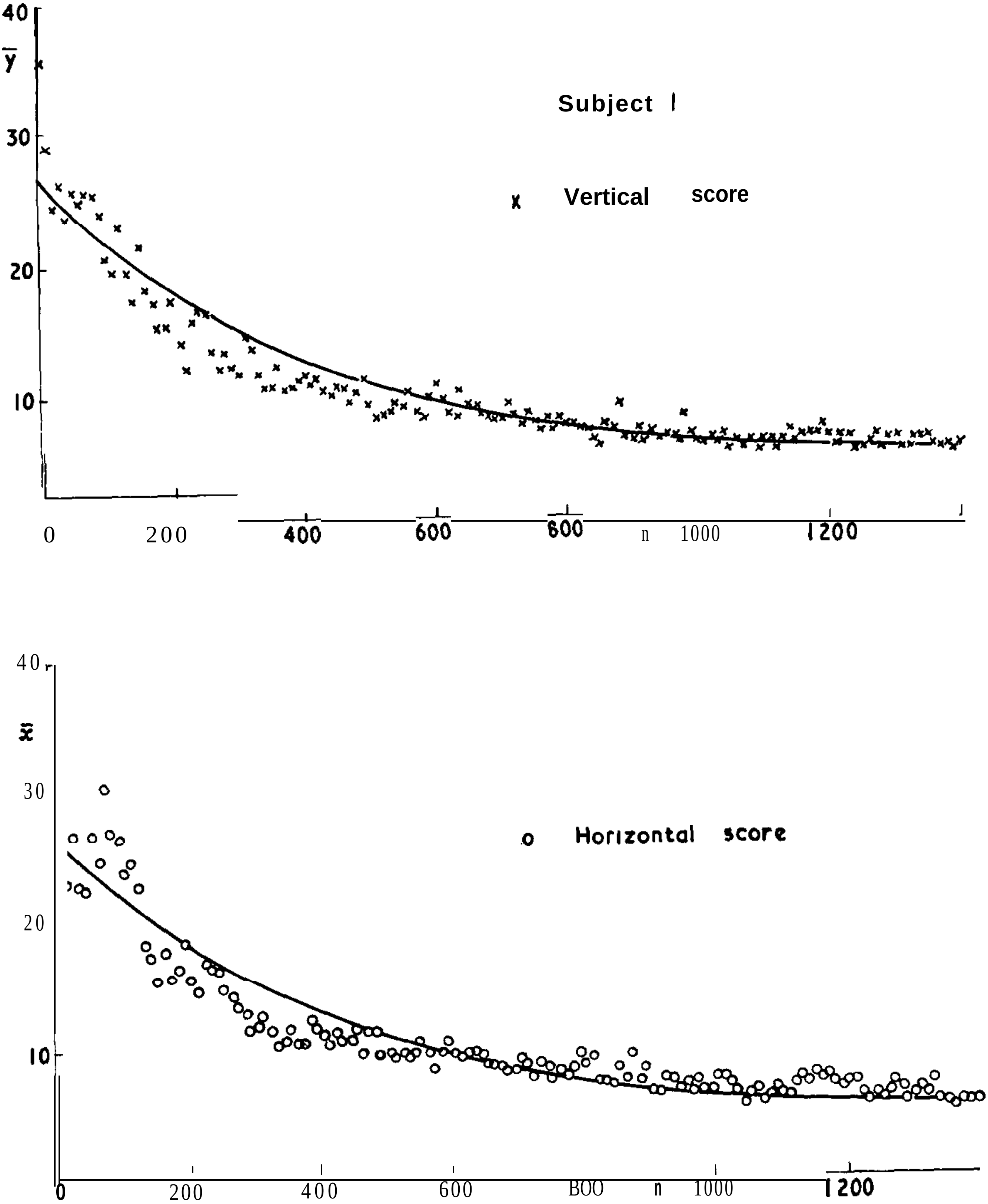
Learning curves. Two -axes position control.

FIG. 6



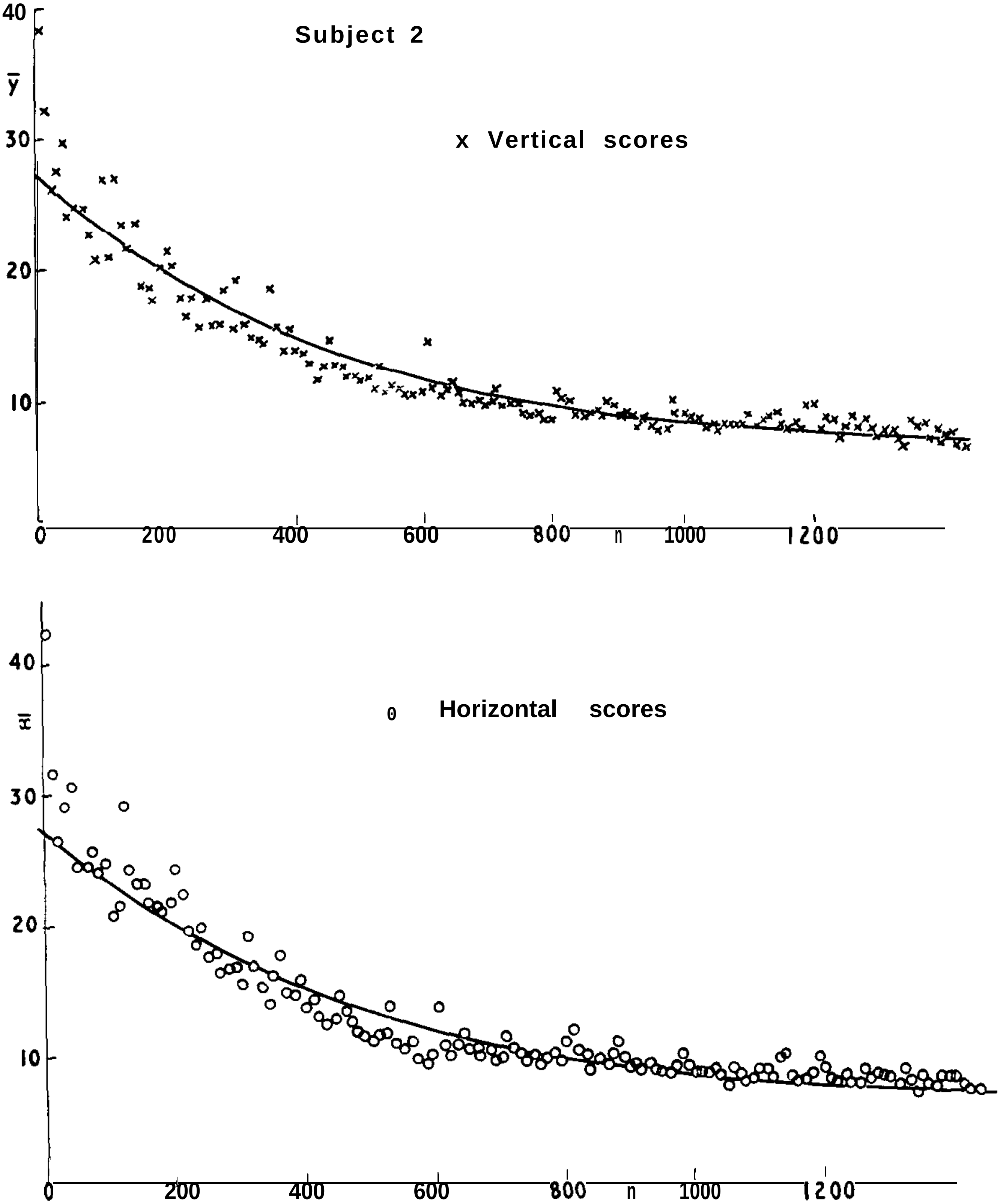
Learning curves. Two -axes velocity control

FIG. 7



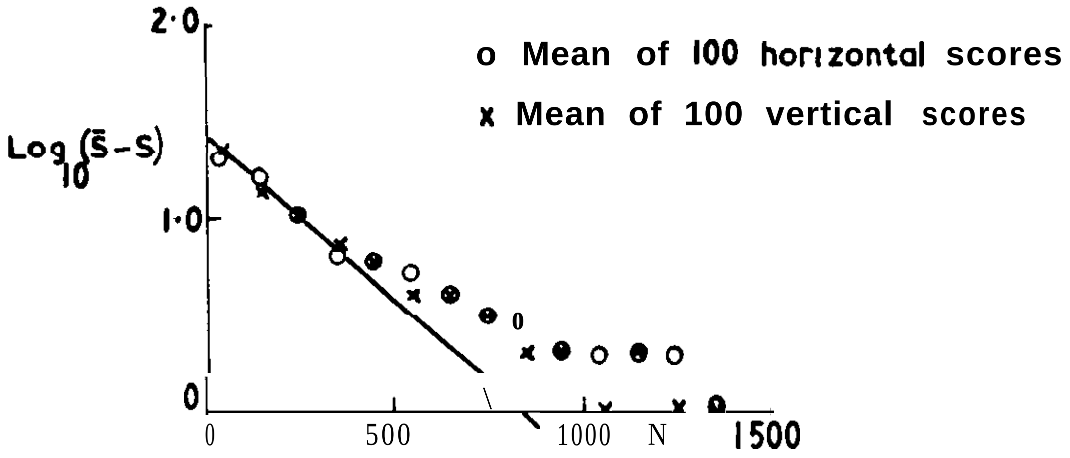
Learning curves Two- axes acceleration control.

FIG. 8



Learning curves. Two-axes acceleration control

FIG.9



Idealised learning curve – Subject 1.
2 -axes acceleration control

A.R.C. C.P.1030
December, 1967
R. C. Hornby and R. Wilson.

THE EFFECTS OF EXTENDED PRACTICE ON PERFORMANCE IN A TRACKING TASK

Experimental measurements of human controller performance have been made during extended periods of practice in visual sine-wave tracking tasks. It has been found that, irrespective of task difficult, RMS error scores decreased to such small magnitudes that differences in scores due to different task variables would have no practical significance. Thus the averaged values of steady scores when tasks are well learnt are meaningless for subject or task difficulty comparisons.

It/

A.R.C. C.P.1030
December, 1967
R. C. Hornby and R. Wilson.

THE EFFECTS OF EXTENDED PRACTICE ON PERFORMANCE IN A TRACKING TASK

Experimental measurements of human controller performance have been made during extended periods of practice in visual sine-wave tracking tasks. It has been found that, irrespective of task difficult, RMS error scores decreased to such small magnitudes that differences in scores due to different task variables would have no practical significance. Thus the averaged values of steady scores when tasks are well learnt are meaningless for subject or task difficulty comparisons.

It/

A.R.C. C.P.1030
December, 1967
R. C. Hornby and R. Wilson.

THE EFFECTS OF EXTENDED PRACTICE ON PERFORMANCE IN A TRACKING TASK

Experimental measurements of human controller performance have been made during extended periods of practice in visual sine-wave tracking tasks. It has been found that, irrespective of task difficult, RMS error scores decreased to such small magnitudes that difference in scores due to different task variables would have not practical significance. Thus the averaged values of steady scores when tasks are well learnt are meaningless for subject or task difficulty comparisons.

It/

It has been shown that the performance scores vary in an exponential manner with the number of task repetitions and it is proposed that en empirical constant related to the rate of decrease of scores be used as a measure of relative task difficulty.

It has been shown that the performance scores vary in an exponential manner with the number of task repetitions end it is proposed that an empirical constant related to the rate of decrease of scores be used as a measure of relative task difficulty.

It has been shown that the performance scores vary in an exponential manner with the number of task repetitions and it is proposed that an empirical constant related to the rate of decrease of scores be used as a measure of relative task difficulty.

C.P. No. 1030

Printed and published by
HER MAJESTY'S STATIONERY OFFICE

To be purchased from
49 High Holborn, London W c 1
13A Castle Street, Edinburgh 2
109 St Mary Street, Cardiff CF1 1JW
Brazennose Street, Manchester M60 8AS
50 Fairfax Street, Bristol BS1 3DE
258 Broad Street, Birmingham 1
7 Linenhall Street, Belfast BT2 8AY
or through any bookseller

Printed in England

C.P. No. 1030

S 0 Code No 23- 9018.30