Vol. XXVI No. 3 PSYCHOLOGICAL REVIEW PUBLICATIONS Whole No. 114 1919

Psychological Monographs

STUDIES FROM THE PSYCHOLOGICAL LABORATORIES OF THE UNIVERSITY OF CHICAGO

THE LEARNING CURVE EQUATION

BY

L. L. THURSTONE, M.E., PH.D.,
Assistant Professor of Psychology, Carnegie Institute of Technology, Pittsburgh, Pa.

PSYCHOLOGICAL REVIEW COMPANY
PRINCETON, N. J.
AND LANCASTER, PA.

AGENTS: G. E. STECHERT & CO., LONDON (2 Star Yard, Carey St., W. C.)
PARIS (76 rue de Rennes)

PRINCETON
UNIVERSITY
PRESS

ACKNOWLEDGMENT

The subjects for this study in learning were students at the Duff Business School in Pittsburgh. I wish to acknowledge the interest and cooperation of Messrs. Spangler and Johnson and of Miss Wilson for making the necessary rearrangements at the Duff School for this investigation. I wish to acknowledge also the advice and interest of Dean J. R. Angell under whose direction this study has been carried out and the many favors of Professor W. V. Bingham without which the work could not have been completed. I am also thankful to Professor Dorweiler of the Carnegie Institute of Technology for his kind advice on some of the mathematical portions of the study.

THE LEARNING CURVE EQUATION

Introduction

INTRODUCTION

The present investigation is essentially an attempt to devise a statistical method for treating learning data. Part I is a discussion of correlation methods and empirical and rational equations. Part II is a description of the learning curve equation and its interpretation. Part III is a discussion of the application of the learning curve equation to typewriter learning. Part IV is a summary.

Learning curves are usually very erratic and for this reason it is necessary to study the general trend of numerous observations instead of the variable individual observations. The methods to be discussed often make it possible to obtain coefficients which express the characteristics of a subject's learning based on all the observations and in such a manner that all observations are as far as possible equally weighted. Quantitative methods in psychology are far in advance of our control over the things measured, and consequently we make ourselves subject to ridicule when refined correlation statistics are applied to measures which are obviously crude. We shall therefore discuss not only the more refined statistical procedure for the learning curve but also some readily applied methods which are adaptable in the study of more or less erratic learning data. Even though refined technique is available we should select the quantitative methods for any particular study so as to keep a fair balance between the certainty of our measures and the statistical niceties by which we treat them.

For the benefit of any reader who wishes to apply the statistical methods to be described for his own learning data I wish to call attention to the first two sections of the summary in which will be found an outline describing the detailed procedure in calculating the learning coefficients.

In applying these methods of learning curve analysis one should be fully aware of their limitations. They are not applicable to the following conditions of learning: 1) when trial and error learning is mixed with generalizations such as in puzzle solving; 2) when the learning is so erratic that it fails to show continuity; 3) when the learning process has not been carried far enough to reveal the nature of the function; which is often the case with apparently linear learning curves; 4) when the learning curve is not plotted in the speed-amount form; 5) when the learning curve fails to show diminishing returns with practice; 6) when the units of formal practice are variable in the different stages of learning (learning measured on different successive scales can not be treated as a continuous function); 7) when the wrong responses are eliminated by ideational learning without giving any objective scores during the process of elimination. Such learning curves have the same appearance as those which contain generalizations.

1) Verbal Statement of Relationship

Our present problem concerns the relationship between practice and attainment in learning. When an observer notes as an element in common experience that attainment increases as practice increases, he may generalize by verbally asserting a positive relation between the two variables. The verbal generalization is so common that it is embodied in what we call common sense. Thus we expect without further verification that twenty hours of practice in a complex function will yield higher attainment than ten hours of practice under roughly similar conditions, but uncontrolled observation does not tell us how much higher.

2) The Correlation Coefficient as an Expression of Relationship

It is possible to express by a single number the degree of relationship between two variables. This is what one attempts to do by a correlation coefficient. The Pearson coefficient of correlation is so derived that when its value is unity the two variables have perfect concomitance. When its value is —1. the two variables have perfect inverse relationship, a rise in one of the variables being always associated with a proportional decrease in the other. A zero correlation establishes the fact of entire absence of relationship within the conditions of the experiment.

A correlation coefficient considerably less than unity may be explained in at least four different ways: 1) the observations themselves may be so inaccurate as to obscure the relationship; 2) the two variables may be related through a common third variable which, if not controlled or kept constant, plays havoc with the experiment; 3) the regression may be non-linear in which case the Pearson coefficient, r, is almost meaningless;* 4) the two variables may be intrinsically independent. Psychological experimentation rarely yields correlations over 0.85 because of the inaccuracy of psychological measures. When a correlation coefficient turns out to be 0.95 or above an empirical equation may properly be substituted for the correlation methods.

In the interpretation of a correlation coefficient one should be careful to note that while a high correlation coefficient does indicate a relation between the variables under the conditions of the experiment, a low coefficient does not indicate the absence of relationship between the variables. The first, second, or third factors enumerated in the preceding paragraph may be responsible for a low coefficient when a high relation really exists. It is perhaps rare that a correlation is calculated with psychological data which is not grossly affected by all three of these factors.

When one variable is immediately contingent on one or more other variables it is advisable to use the method of partial correlation to establish the relation. Thus the volume of a box is contingent on the three variables, length, width, and depth. Now, if a heterogeneous collection of wooden boxes were to be measured as to all four of these attributes and the correlation coefficient between volume and length determined, it would undoubtedly turn out to be positive and significant because long boxes are usually more voluminous than short boxes. But the

*The term *regression* was introduced by Galton in connection with his statistical studies in the heredity of stature. It is the equation of the best fitting line for a series of paired observations of two variables. The use of the term seems to be restricted to data of considerable dispersion and is not used for those observations in the exact sciences which do not involve serious scatter. The term *linear regression* refers to the equation of a regression line which is straight, as contrasted with *non-linear regressions* which are curved. See Yule, "Textbook in Theory of Statistics," p. 176 and references 2 and 3 on p. 188.

coefficient would not be unity because of the two other variables which were left out of consideration. This would illustrate case 2 in the preceding paragraph. We may distinguish two methods of handling this type of relation. a) We may control the extraneous variables by keeping them constant while measuring the two variables with which we are immediately concerned. This is the customary procedure in the physical sciences. Thus when verifying Boyle's Law we keep the temperature of the gas constant, but when verifying Charles' Law we keep the pressure constant. Except for these precautions neither of the two laws would be observed. In the biological sciences we do not have such ready control over the extraneous variables and the best we can do is to measure them also, while allowing them to vary at will. b) In these cases we use a second method, namely, partial correlation. Thus if we calculate the partial correlation coefficient between volume and length of the collection of wooden boxes, with the depth and width accounted for, we obtain a higher coefficient than when the two extraneous variables were ignored. The advantage of the partial correlation method is that it enables us to control the extraneous variables analytically without having any physical control over them.

There are some limitations of the correlation methods which every experimenter should keep in mind in order to guard against erroneous conclusions. One of the limitations of the partial correlation method is that it assumes the combined effect of the several independent variables on the dependent variable to be additive, a condition which rarely obtains. This limitation is much more serious than the inadequacy of the correlation coefficient for non-linear regressions. A non-linear regression can usually be rectified by one of the algebraic artifices used in connection with empirical equations, but the assumption that the several independent variables produce their effect on the dependent variable in an additive manner can not be handled by any predetermined statistical method. Thus, returning to the box illustration, the best measure we could obtain by the method of partial correlation is expressed in the form

$$v = k_1 d + k_2 w + k_3 l. \qquad 1$$

But the true formula for the volume takes the form

$$v = k.d.w.l. \qquad 2$$

Now, this type of relation is not revealed by the partial correlation method, nor is the method adequate for any of the thousands of ways in which the several variables may combine except the additive one.

If we are not content with merely stating in quantitative form the degree of relationship between the two variables but wish to formulate a method of prediction, we use the regression equation. This equation is derived from the Pearson coefficient which merely states in numerical form the *degree* of relationship between the two variables. It places the information at our immediate command for the purpose of prediction. Given one of the unknowns, the other can be found either from the regression lines or from the regression equation which is simply an algebraic description of the regression line.

It is quite conceivable that a low value of attribute A may be associated with either high or low value of attribute B, whereas a high value of attribute A may be associated with only high values of attribute B. Similarly a low value of attribute B may be associated with low values of attribute A only, whereas high values of B are associated with either high or low values of attribute A. Whenever such conditions obtain the Pearson correlation coefficient is inadequate to express the complete relationship. In these cases, called non-linear regressions, it is advisable to calculate another kind of coefficient which is called the eta coefficient, η,* or correlation ratio. The significance of the correlation ratio may perhaps be made more apparent by the analogue *All dogs are quadrupeds but all quadrupeds are not dogs.* What would correspond to the correlation ratio of dogs on quadrupeds would be very high because all dogs are quadrupeds. But the correlation ratio of quadrupeds on dogs would be low, for only a few of the quadrupeds are dogs. The Pearson correlation coefficient for a relation of this type would be positive but low.

*For a brief statement giving the derivation of the correlation ratio, see Yule, p. 204.

There are a number of algebraic artifices by means of which a non-linear regression may be rectified. The value of such devices becomes apparent when it is considered that such otherwise exceedingly useful tools as are available in the correlation methods are inapplicable as long as the regressions are non-linear. The investigator must rely on his own ingenuity in rectifying a non-linear regression. Some of these methods will be considered in connection with the learning curve equation.

3) Empirical Equations

Every equation can be represented by a line in a diagram and practically every line encountered in quantitative experimental work can be represented by an equation. Thus the regression equation is only an algebraic way of describing the regression line of a scatter diagram, or, putting it the other way, the regression line is a graphical description of the regression equation. Each tells the same story in its respective language.

An empirical equation is an equation selected to fit a given set of data. The observations give us the diagram and if we find an equation whose line coincides with the general trend of the observations, it may be used interchangeably with the diagram for predicting one of the attributes when the other is given. When the observations indicate a linear relation we can derive the corresponding equation with very little trouble, but when the observations fall along a curve and when they are badly scattered the finding of the most representative empirical equation sometimes taxes the investigator's ingenuity. I shall describe the routine steps in determining the empirical equation for a linear relation by means of an example and will show that it turns out to be identical with the regression equation for the same data.

Figure 1 is a diagram of the relation between two hypothetical variables X and Y. Each of the small circles represents a hypothetical observation; the solid line represents the general trend of the observations. This line may be used for the purpose of prediction. Thus if we know that on a certain occasion attribute X had a numerical value of 9, the attribute Y must have been very close to the value 63, as read from the chart. Our problem now is to describe this line algebraically so that the prediction may be made by means of a formula instead of by the diagram.

a) *Method of inspection.* This procedure is the simplest but it can only be applied when the relation is close, as it is in the present illustration. We first indicate the observations by small circles or dots on the diagram. Then we draw by inspection the best fitting straight line through the general trend of the observations. The equation of a straight line always takes the form

$$Y = a + b.X \qquad 3$$

The y-intercept is 30.8 and it is the constant *a*. The slope of the line is 3.56 and it is the constant *b*. Hence the equation for the line is

$$Y = 30.8 + 3.56X \qquad 4$$

This equation may be used interchangeably with the diagram in predicting one of the attributes when the other is known. The procedure is so simple and direct that it would be universally used, were it not for the fact that when the observations scatter badly, it is difficult to draw the best fitting straight line by inspection. Moreover, the same line can not properly be used in predicting X from Y as in predicting Y from X when the data are scattered. In these cases we have recourse to two other types of procedure, namely, the method of the regression equation and the method of least squares.

b) *Method of regression equation.* In figure 2 we have represented a hypothetical set of data which are quite scattered. By the usual correlation methods we obtain the following constants:

$r = +\ 0.59$
$\sigma_x = 4.58$
$\sigma_y = 3.03$
$n = 50.$
$m_x = 10.$
$m_y = 8.$

The regression equation with two variables for predicting X from Y takes the form

$$x = r_{x.y} \frac{\sigma_x}{\sigma_y} y$$

in which x and y are deviations of X and Y from their respective means. Rewriting this equation in terms of the variables X and Y instead of in terms of the deviations from their means, we have

$$X - m_x = r_{xy} \frac{\sigma_x}{\sigma_y} (Y - m_y) \qquad 6$$

in which m_x and m_y are the arithmetic means of X and Y respectively.

Substituting the numerical values into equation 6 we have

$$X - 10 = \frac{0.59 \times 4.58}{3.03} (Y - 8) \qquad 7$$

and simplifying, we obtain

$$X = 0.89Y + 2.88 \qquad 8$$

This is the regression equation ready for use. By means of it we predict X when the value of Y is known. When the data are not seriously scattered it is safe to fit the line by inspection and determine the empirical equation by the shorter method, but with the data of figure 2 the regression lines can hardly be judged by inspection.

By analogy the regression equation for predicting Y from X takes the form

$$y = r_{xy} \frac{\sigma_y}{\sigma_x} x \qquad 9$$

which when stated in terms of the variables instead of in terms of the deviations from their respective means becomes

$$Y - m_y = r_{xy} \frac{\sigma_y}{\sigma_x} (X - m_x) \qquad 10$$

Substituting and simplifying, as before, we get as a numerical statement of the relation that

$$Y = 0.39X + 4.1 \qquad 11$$

which is ready for use in predicting the most probable value of Y for a known value of X.

These regression equations may be obtained with less arithmetical labor, particularly if only one of the regression equations is needed, by the method of least squares.

c) *Method of least squares.*[*] The method of least squares is an aid in finding the best fitting straight lines for representing a series of observations. It can be applied also to curves but that often leads to awkward mathematical maneuvering. The method of least squares gives a line such that the sum of the squares of the deviations of the independent variable from the regression line is a minimum. It is the best fitting straight line for the observations. The method gives the numerical values of the constants *a* and *b* in the equation

$$X = a + b \cdot Y \qquad 3$$

which represents any straight line and by means of which we can predict the most probable value of X from a given value of Y. The application of the method consists in solving the two following formulae:

$$a = \frac{\Sigma(Y) \cdot \Sigma(X \cdot Y) - \Sigma(Y^2) \cdot \Sigma(X)}{[\Sigma(Y)]^2 - n \cdot \Sigma(Y^2)} \qquad 12$$

$$b = \frac{\Sigma(Y) \cdot \Sigma(X) - n \cdot \Sigma(X \cdot Y)}{[\Sigma(Y)]^2 - n \cdot \Sigma(Y^2)} \qquad 13$$

Substituting the appropriate sums from the data we find that $a = 2.9$ and $b = 0.89$. Hence the equation of the best fitting line for predicting X from Y is

$$X = 2.90 + 0.89Y \qquad 14$$

It should be noted that this equation, as determined by the method of least squares, is identical with the regression equation 8 which was determined by correlation methods.

By analogy, the equation for the straight line by which the most probable value of Y may be determined from a known value of X is

$$Y = c + d \cdot X \qquad 15$$

The constants *c* and *d* may be determined from the original observations by the formulae:

$$c = \frac{\Sigma(X) \cdot \Sigma(X \cdot Y) - \Sigma(X^2) \cdot \Sigma(Y)}{[\Sigma(X)]^2 - n \cdot \Sigma(X^2)} \qquad 15^a$$

$$d = \frac{\Sigma(X) \cdot \Sigma(Y) - n \cdot \Sigma(X \cdot Y)}{[\Sigma(X)]^2 - n \cdot \Sigma(X^2)} \qquad 16$$

Substituting the numerical value of the data from figure 2 into equations 15 and 16 we obtain: $c = 4.09$, and $b = 0.39$. Substituting these numerical values into 15 we get

$$Y = 4.09 + 0.39X \qquad 17$$

by means of which we may predict the most probable value of Y from a known value of X.

It can be shown readily that when the regression equations have been determined by the method of least squares the Pearson coefficient of correlation is expressed by the relation

$$r = \sqrt{b \cdot d} \; * \qquad 18$$

in which b and d are identical with the regression coefficients.

The above calculations make it apparent that the method of least squares gives us a pair of regression lines which are identical with those obtained by the correlation methods. For the purpose of stating quantitatively the degree of relationship between two variables it is desirable to calculate the correlation coefficient. When the regression equation is of primary interest it can be calculated to advantage by the method of least squares particularly if only one of the regressions is needed.

4) Rational Equations

A rational equation is derived from known relations and is verified by experimental observation. From a philosophic point of view one may argue that all equations used by science are in the last analysis empirical but in practice there is a far cry between fitting an empirical equation to a series of observations and the ability to predict the observed relations on the basis of a rational equation.

Psychology has very few bona fide *laws* on which we can build a system of quantitative prediction and control. Thus practically all we can do with the problem of learning is to observe the function and describe it. The present attempt is to describe it quantitatively by an empirical equation. Some day we shall possess in psychology a coordinated system of really workable concepts with objective reference by which we may be able to predict and control at least certain aspects of behavior by rational equations or their equivalents.

*See Yule, p. 203.

II. THE LEARNING CURVE EQUATION

1) Purpose of the Equation

When the learning function for a simple coordination proceeds undisturbed by external or internal distraction it usually follows a law of diminishing returns. In the majority of learning curves the amount of attainment gained per unit of practice decreases as practice increases. Exceptions to this tendency are found in studying the learning of complex processes such as a foreign language, and when successive generalizations are involved such as puzzle solving and the like. These exceptions sometimes take the form of a positive acceleration at the initial stage of the learning, plateaus during the course of learning and erratic advance of attainment. But these irregularities should not stand in the way of an attempt to express the learning function as a law provided that we do it with due conservatism in its interpretation. All we can hope to do in thus expressing the learning function is to formulate what can with considerable certainty be considered as the typical relation between practice and attainment.

Besides giving the satisfaction of formulating the relation between practice and attainment, the use of an equation for this relation enables one to predict the limit of practice before it has been attained, provided that the learning follows the law of diminishing returns. It also enables one to differentiate for various purposes the rate of learning from the limit of practice since these two attributes are undoubtedly independent. It enables us to state how much preceding practice the subject has experienced under the assumption that the learning function followed the same law before and after the formal measurements. Another use for which the equation can be of service is in the analysis of the relation between the variability in learning and other mental attributes. The problems of formal discipline may be investigated by ascertaining whether a succession of learning processes, all of the same type, yields any rise in the limit of practice, or a higher rate of learning, or a greater consistency of learning in the successive learning processes. Some of these coefficients may be more susceptible than others to modification by successive repetition of the same type of learning. This would in reality be studying the problems of learning how to learn. All questions of transfer of training may be investigated by the learning equation and the transfer effect may be differentiated into its psychological components. Thus, continued practice in learning poetry may show no rise of the practice limit, but a considerable rise in the rate at which that limit is approached and in a decrease of the variability of the learning. Relearning may be found to approach the same limit of practice as the initial learning but it may proceed at a higher rate, and this rate can be stated as a coefficient which is independent of the amount of previous practice in each learning process. The laws of forgetting are expressible in terms quite similar to those here used for the learning function. It is not at all unlikely that these coefficients may come to be significant in individual psychology quite apart from their immediate utility as descriptive attributes of the learning function. The preceding remarks have, I hope, justified my attempt to devise a method for investigating the learning, memory, and forgetting functions.

2) The Equation

After experimenting with some forty different equations on published learning curves I have selected a form of the hyperbola as being for practical purposes the most available. It takes the form

$$Y = \frac{L \cdot X}{X + R} \qquad 19$$

in which

Y = *attainment* in terms of the number of successful acts per unit time.

X = *formal practice* in terms of the total number of practice acts since the beginning of formal practice.

L = *Limit of practice* in terms of attainment units.

R = *Rate of learning* which indicates the relative rapidity with which the limit of practice is being approached. It is numerically high for a low rate of approach and numerically low for a high rate of approach.

Equation 19 represents a learning curve which passes through the origin, i.e., it starts with a zero score at zero formal practice. The majority of learning curves start with some finite score even at the initial performance. For learning curves which do not pass through the origin, the equation becomes

$$Y = \frac{L(X+P)}{(X+P)+R} \tag{20}$$

in which P = *equivalent previous practice* in terms of formal practice units.

Figure 3 represents the learning curve for subject No. 23 in the group of fifty-one typewriter students to be discussed in a later section. This curve is plotted between attainment, Y, in terms of the number of words written in a four minute test given weekly for seven months, and formal practice (X) in terms of the total number of pages written since entering the course. We shall call this type of curve the *speed-amount* curve to distinguish it from other ways of plotting the same data.

Equation 19 may be rectified as follows:

$$Y = \frac{L \cdot X}{X + R}$$

$$XY + R \cdot Y = LX$$

$$X + R = L\left(\frac{X}{Y}\right) \tag{21}$$

This equation is linear if X/Y is plotted against X. Similarly equation 21 may be rectified when written in the form

$$X + (R + P) = L\frac{(X + P)}{Y} \tag{22}$$

which becomes linear when $(X + P)/Y$ is plotted against X.

When so rectified, the constants L, R, and P may be determined by several different methods, the choice between which depends on the scatter of the data, the desired accuracy, and the number of curves one has to calculate. We shall describe four methods of calculating the coefficients.

a) *Method of least squares*

Case 1: when the learning curve passes through the origin: Arrange the data as in table 1. Calculate X/Y and tabulate. Plot X/Y against X as in figure 4. For convenience we shall call

$$\frac{X}{Y} = Z \tag{23}$$

The reader will notice that the learning data as plotted in figure 4 falls practically in a straight line whereas the same data in figure 3 takes the typical learning curve form.

The equation for the best fitting straight line of figure 4 can be represented by the equation

$$Z = c + d \cdot X \tag{24}$$

The constants *c* and *d* are determined from the table of data by the formulae:

$$c = \frac{\Sigma(X) \cdot \Sigma(X \cdot Z) - \Sigma(X^2) \cdot \Sigma(Z)}{[\Sigma(X)]^2 - n \cdot \Sigma(X^2)} \tag{25}$$

$$d = \frac{\Sigma(X) \cdot \Sigma(Z) - n \cdot \Sigma(X \cdot Z)}{[\Sigma(X)]^2 - n \cdot \Sigma(X^2)} \tag{26}$$

which are simply the least square formulae (15) and (16) rewritten for X and Z. Substituting the proper sums, we have $c = 0.42$, and $d = 0.0041$. Hence the equation for X and Z becomes

$$Z = 0.42 + 0.0041X \tag{27}$$

which by replacing X/Y for Z and transposing becomes

$$Y = \frac{244X}{X + 102} \tag{28}$$

in which the predicted limit of practice, L, is 244 words in four minutes, and the rate of learning, R, is 102. The constants L and R may also be determined by the relations

$$R = c/d$$
$$Z = 1/d$$

Plotting equation (28) we obtain the solid line in figure 3. It will be noticed that this curve fits quite well the general trend of the observations which are indicated by the small circles. This method of stating algebraically the relation between practice and attainment is of course not applicable unless the speed-amount curve for the data takes the typical hyperbolic form.

Case 2: when the learning curve does not pass through the origin. When the learning curve does not pass through the origin it can be rectified by slightly different procedure. We shall take as an illustration the combined curve for a group of fifty-one subjects studying typewriting. Figure 5 represents the average speed of typewriting against the total number of pages written since entering the course. It is seen to be a fairly smooth and regular curve.

The equation for the learning curve which does not pass through the origin is

$$Y = \frac{L(X + P)}{(X + P) + R} \qquad 20$$

or, if we call

$$P + R = K \qquad 32$$

for convenience, we have, instead of equation (20)

$$Y = \frac{L(X + P)}{X + K} \qquad 33$$

This equation can be rectified as follows: When $X = 0$, and $Y = Y_1$, Y_1 being the initial attainment score,

$$Y_1 = \frac{L \cdot P}{K} \qquad 34$$

and hence equation (33) becomes

$$\frac{X \cdot Y}{Y - Y_1} = L \frac{X}{Y - Y_1} - K \qquad 35$$

This equation is linear if $XY/(Y-Y_1)$ is plotted against $X/(Y-Y_1)$, in which case L is the multiplying constant and K is the additive constant.

Plotting the data represented in figure 5 in this manner we obtain figure 6 in which the learning data appear as a straight line. This line may be represented by the equation

$$S = a + b \cdot T \qquad 36$$

in which

$$T = \frac{X}{Y - Y_1} \text{ and } S = \frac{X \cdot Y}{Y - Y_1}$$

The numerical values of a and b may be determined by the following least square formulae which are identical with equations *12* and *13*, except for the analogous notation.

$$a = \frac{\Sigma(T)\cdot\Sigma(S\cdot T) - \Sigma(T^2)\cdot\Sigma(S)}{[\Sigma(T)]^2 - n\cdot\Sigma(T^2)} \qquad 37$$

$$b = \frac{\Sigma(T)\cdot\Sigma(S) - n\cdot\Sigma(S\cdot T)}{[\Sigma(T)]^2 - n\cdot\Sigma(T^2)} \qquad 38$$

Substituting the proper sums we find that $a = -148.$ and $b = 216.$ Hence

$$S = 148. + 216.T \qquad 39$$

which is the equation of the solid line in figure 6. This equation may be transposed into the original form of equation *20*, or we may write it in that form directly by the following relations:

$$a = K$$
$$b = L$$
$$P = \frac{a\cdot Y_1}{L}$$
$$R = K - P$$

All of the constants K, L, Y, P, and R, are positive when applied to learning curves. It should be noted that Y_1 is a representative original score determined by projecting the learning curve back to the y-axis. In figure 5 the actually observed initial score was used since it is continuous with the rest of the data. But it is occasionally necessary to select a representative initial score since Y_1 is weighted in this procedure more than any of the other points. The numerical values of these constants for the data of figure 5 are as follows:

$$L = 216.$$
$$P = 19.$$
$$R = 133.$$

Substituting these constants in equation *33* we have

$$Y = \frac{216.(X + 19.)}{X + 148.} \qquad 40$$

which when plotted becomes the solid line of figure 5. The reader will notice that this equation, as represented by the solid line in figure 5, is a beautiful fit for the data, and it justifies our use of equations *19* and *20* to represent the hyperbolic form of learning curve.

The predicted limit of practice L which is 216. words in four minutes, is of course based on the assumption that the learning curve would continue as uniformly beyond the measurements as it did during the measurements. This limitation must be kept in mind and we shall therefore differentiate between the predicted limit and a limit of practice which has been practically attained. The equivalent previous practice (P) is *19* pages which we may interpret as the average number of pages of typewriting to which the previous general experience of our subjects was equivalent. This interpretation of the constant P is also limited by the assumption that the unmeasured learning function followed the law which the measurements reveal. One circumstance which bears out this assumption is that those learning curves which actually do pass through the origin and which do not show positive acceleration usually follow this curve law when the coordinates are properly chosen. The curve of figure 5 does not pass through the origin but this is explainable by the fact that a person who has never touched a typewriter will in four minutes make some finite score even though handicapped by using *the hunt and punch method.*

b) *Method of inspection*

When the observations fall very nearly in a straight line as they do in figure 6 it is hardly necessary to plough through the arithmetical labor involved in evaluating the constants *a* and *b* of equation *36* by the method of least squares unless one has ready access to a calculating machine. After plotting figure 6 one may draw at sight the best fitting straight line through the general trend of the data and evaluate the constants from any of the following relations:

$$\text{y-intercept} = a = K$$
$$\text{x-intercept} = K/Y_1$$
$$\text{slope} = L$$
$$P = \frac{a \cdot Y_1}{L}$$
$$R = K - P$$

By this graphical procedure much labor is saved in calculating the learning curve constants and the method is identical with the preceding in principle.

c) *Method of three equidistant points*

The learning coefficients may be determined from three selected points with less labor than when all the observations are taken into account. These three points should be so selected that they represent the general trend of the learning curve.

Let the three selected points be denoted X_1Y_1; X_2Y_2; and X_3Y_3. Let X_1 be zero, X_3 the total amount of practice and X_2 the midpoint between X_1 and X_3. Let the Y-values be the most representative ordinates to the curve. Then

$$X_3 = 2 \cdot X_2 \qquad 41$$

By substituting these values into equation *20*, transposing and simplifying, we obtain

$$K = \frac{X_3(Y_2 - X_3)}{Y_3 + Y_1 - 2Y_2} \qquad 42$$

$$L = \frac{Y_3(X_3 + K) - Y_1 \cdot K}{X_3} \qquad 43$$

$$P = \frac{Y_1 \cdot K}{L} \qquad 44$$

$$R = K - P \qquad 45$$

From these relations we may determine the numerical values of the learning coefficients in terms of the three equidistant points.

When the curve passes through the origin both X_1 and Y_1 are zero. The coefficients may then be determined by the following somewhat simpler relations:

$$K = \frac{X_3(Y_2 - Y_3)}{Y_3 - 2 \cdot Y_2} \qquad 46$$

$$L = \frac{Y_3(X_3 + K)}{X_3} \qquad 47$$

$$R = K$$

P is zero because when the initial score is zero the equivalent previous practice is zero.

3) Interpretation of Learning Constants

We have seen that the learning curve equation *20* fits very well the learning data to which we have applied it. In order to bring out the interpretation of the learning coefficients we shall compare several learning curves with high and low numerical values of the coefficients.

In figure 7 we have two hypothetical learning curves with different physiological limits but with identical rates of approach. Figure 8 represents two hypothetical learning curves, both approaching the same limit of practice, one at a high rate and the other at a low rate. Figure 9 represents two hypothetical learning curves with same limit of practice, and with the same rate of approach, but differing in the amount of previous practice. Curve A represents forty units of previous practice while curve B represents no previous practice. The two curves are identical in shape, the only difference between them being that curve B is forty x-units to the right of curve A. The same interpretation would be reached if the two curves were superimposed and the formal practice measurements started at the origin for curve B and after forty practice units for A.

4) The Coordinates for Learning Curves

So far we have considered learning curves plotted only between the coordinates X (total number of practice acts since the beginning of practice) and Y (the number of successful acts per unit time). Learning curves have, however, been plotted with various units for the coordinates and we shall consider several of these together with some inferences that may be drawn from the translation of learning data from one system of units to another.

The speed-amount curve is the name we shall use to designate the form of learning curve we have been considering. It is plotted as speed, Y, against amount of practice, X. It may be represented by equations 19 and 20 when it reveals the typical hyperbolic form.

The time-amount curve is plotted as time, t, per unit amount of work against total amount of work, X, since the beginning of practice. It is evident that the ordinates of this type of curve will be proportional to the reciprocals of the speed-amount curve for the same data. Hence we may define *t* as

$$t = \frac{C}{Y} \qquad 48$$

where *t* is the time per unit amount of work and *C* is a constant. Limiting ourselves to the curves of diminishing returns we have, as the equation of the time-amount curve

$$t = \frac{C(X + K)}{L(X + P)} \qquad 49$$

The constant *C* is only significant in translating learning curves from one form to the other. Applying the equation directly to learning data the constant *C* may be dropped. In that case

$$t = \frac{X + R}{L(X + P)} \qquad 50$$

This equation may be rectified by the procedure previously outlined for equations *19* and *20*. When P = zero, we have

$$t = \frac{X + R}{L \cdot X} \qquad 51$$

which can be rectified by plotting tX against X.

In order to determine whether equations *50* and *51* really fit the time-amount curve throughout its range I have given a long substitution test to one of my students. He took the test seventeen times, once a day, and reached what is for all practical purposes a practice limit. The time-amount curve for this learning test is represented in figure 11. In figure 12 I have rectified the data by plotting the products tX against X. The reader will notice that the speed-amount curve is hyperbolic. It is quite gratifying that the learning records for an individual subject follow the hyperbolic law so closely. In order to avoid erratic scores from individual subjects it is absolutely essential that they work under uniform conditions with a minimum amount of distraction. The student whose substitution learning is represented in figures 11 and 12 took the test once a day only and always at 1 P. M. The test consisted in making six hundred substitutions at each sitting.

The time-time curve is the learning curve plotted between the time, t, per unit amount of work and the total time, T, devoted to practice. An empirical equation may be derived for this type of curve from the assumed hyperbolic form of the speed-amount curve. The total time is the summation $\Sigma t \cdot dx$ for the whole period of learning. Hence

$$T = \int t \cdot dx \qquad 52$$

But from equation *50*

$$t = \frac{X + K}{L(X + P)} \qquad 50$$

and hence

$$T = \int \frac{X + K}{L(X + P)} dX$$

which may be written

$$T = \frac{1}{L} \int dX + \frac{K - P}{L} \int \frac{dX}{X + P}$$

Integrating, we obtain

$$T = \frac{X}{L} + \frac{K - P}{L} \log (X + P) + C_1 \qquad 53$$

which gives the equivalent total time T in terms of X. Stating X explicitly from equation *50* and substituting in equation *53* gives the desired relation between T and t as

$$T = \frac{K - t \cdot L \cdot P}{(L \cdot t - 1)L} + \frac{K - P}{L} \log \left(\frac{K - P}{L \cdot t - 1}\right) + C_2 \qquad 54$$

While this equation does give us a relation between T and t as derived from the hyperbolic speed-amount curve it is too unwieldy to be practically feasible. We are hardly justified in using so complex an empirical equation for learning data.

The speed-time curve is plotted between the speed Y (number of successful acts per unit time) and the total amount of time, T, devoted to practice. An equation between Y and T may be derived by stating X explicitly from equation *20* in terms of Y and substituting this for X in equation *53* which gives

$$T = \frac{P - K}{Y - L} + \frac{K - P}{L} \log Y - \frac{K - P}{L} \log (L - Y) + C_2 \qquad 55$$

While this equation is too cumbersome for extensive use it serves one very interesting function in that it sheds light on the question of positive acceleration in learning curves.

5) INITIAL POSITIVE ACCELERATION IN THE SPEED-TIME CURVE

Equation 55 represents the speed-time curve. It may be simplified by letting

$$A = P - K \text{ and } B = \frac{K - P}{L} \qquad 55a$$

when it becomes

$$T = \frac{A}{Y - L} + B \cdot \log Y - B \cdot \log (L - Y) + C_2 \qquad 56$$

The first derivative with respect to Y is

$$\frac{dT}{dY} = -\frac{A}{(Y - L)^2} + \frac{B}{Y} + \frac{B}{L - Y} \qquad 57$$

The second derivative is

$$\frac{d^2T}{dY^2} = \frac{2A}{(Y - L)^3} - \frac{B}{Y^2} + \frac{B}{(L - Y)^2} \qquad 58$$

which when simplified becomes

$$\frac{d^2T}{dY^2} = \frac{B \cdot (L)^2 \cdot (3 \cdot Y - L)}{Y^2 \cdot (L - Y)^3} \qquad 59$$

since $A = -BL$ from equation 55a

Equating the second derivative to zero, we have:

$$\frac{B \cdot (L)^2 \cdot (3 \cdot Y - L)}{Y^2 \cdot (L - Y)^3} = 0$$

which is true when Y has the value $L/3$.

This demonstrates the presence of a point of inflection in equation 55 at the value $L/3$ for Y. The psychological significance of this relation may be stated as follows:

The learning curve in the speed-time form must necessarily have an initial positive acceleration which changes to a negative acceleration when the attainment has reached one-third of the limit of practice. This conclusion is contingent on the assump-

tion that the learning curve in the speed-amount form is hyperbolic, an assumption which has been empirically shown to be safe for the majority of learning curves. As has already been said, the speed-amount curve is usually hyperbolic but not always. These assertions regarding the speed-time curve are not applicable when the speed-amount curve for the same data is not hyperbolic. The positive acceleration can not, of course, be obtained when the initial score is greater than one-third of the practice limit. It can only be observed when the initial score is less than one-third of the practice limit.

In order to test empirically the above finding with regard to initial positive acceleration I have plotted in figure 13 the average typewriting speed for fifty-one subjects against weeks of practice (the speed-time curve) instead of against total number of pages written (the speed-amount curve). The average practice limit for this group has already been found to be 216 words in four minutes according to the speed-amount curve for the same data. The reader will notice the initial positive acceleration followed by negative acceleration, and also that the point of transition from positive to negative acceleration takes place at a writing speed of about seventy words in four minutes, as it should do according to our analysis of the speed-time curve. If this finding will stand the test of further experimentation it is obviously of considerable diagnostic value for the psychologist who can by means of it predict the practice limit when attainment reaches one-third of its limit. The limitation in the use of this relation is mainly in the erratic improvement in complex coordinations which are learned under variable conditions of distraction and in the occasional deviations from the typical hyperbolic form of the speed-amount curve.

6) Other Possible Equations

Before closing the discussion on the learning curve equation as such it might not be out of place to mention a few of the other equations which I have tried to use for learning data. These will not be of interest to the general reader but may be of interest to those who wish to try their hand at other empirical equations for the learning function.

One of these equations for the speed-amount curve is

$$Y = L[1 - \frac{1}{e^{ax}}] \qquad 60$$

where e is the Naperian base or some other constant. This equation can not readily be rectified except by trying successive values for L. When the proper value for L is found it can be rectified when written in the form

$$\log (L - Y) = \log L - a \cdot X \cdot \log e \qquad 61$$

by plotting $\log(L - Y)$ against X. If the curve does not pass through the origin equation *60* becomes

$$Y = L[1 - \frac{1}{e^{a(x+p)}}] \qquad 62$$

which is rectified if the proper numerical values of L and P are found by writing it in the form

$$\log(L - Y) = \log L - a(X+P)\log e \qquad 63$$

and plotting $\log(L - Y)$ against $(X+P)$. This equation gives a fair approximation to the speed-amount curve but it does not fit nearly as well as the hyperbolic form previously considered. It can be rectified graphically by plotting Y-increments against X but this procedure is not feasible unless the individual observations are more consistent than they usually are for learning data. The constants L and P can also be determined graphically from three selected points. If X_1Y_1, X_2Y_2, and X_3Y_3 be three points on the curve, equidistant on the axis of abscissae, then the two lines X_3Y_2; X_2Y_1 and X_3Y_3; X_2Y_2 will intersect in a point which is on the asymptote parallel to the axis of abscissae, thus determining the constant L graphically. This equation gives a fair approximation to the speed-amount curve but it does not fit nearly as well as the hyperbolic form previously considered.

Another equation which gives a fair approximation to the learning curve is

$$Y = L(B^{-\frac{C}{X}}) \qquad 64$$

in which B and C are constants. It can be rectified by writing it in the form

$$\log Y = \log L - \frac{C}{X} \log B \qquad 65$$

and plotting logY against $1/X$. It has the advantage of simplicity and it can be used to represent an initial positive acceleration. But as far as I have been able to determine the constant L does not agree as well with observed values as the hyperbolic form.

One could perhaps write an indefinite number of exponential, trigonometric and other functions to represent the learning curve but as long as the simple equation *20* with its various transformations fits the data, and as long as we do not have the basis for a rational equation for learning I have been content to abide by it.

1) The Subjects

Eighty-three students at the Duff Business School in Pittsburgh took one four minute typewriter test once a week during the school year 1916-17. The tests were begun in September and continued until the middle of April. The subjects practiced two hours of school schedule time every day, five days a week. No tests are available for the first three weeks of practice because teachers of typewriting who use the touch system prefer not to give tests from straight copy until the mechanism of the typewriter and the key board have been mastered. This takes from three to seven weeks, depending on the maturity, adaptability and industry of the students. Practically all of the subjects had finished the grammar school, a number of them had completed one or two years of high school, and several had finished a four year high school course. Their average age was about seventeen years. In order to obtain an initial typewriting score I asked ten of my students who had never touched a typewriter to take a four-minute test. The average score for this group was 27 words in four minutes and this is used with the other data as an average initial score in typewriting.

Of the eighty-three subjects who took the tests thirty-two were eliminated, leaving fifty-one subjects for the major study. The causes of elimination are indicated in the following table:

Original size of group.........		83
Irregular attendance	20	
Unusually irregular performance	3	
Apparent linearity of learning curve	5	
Delayed positive acceleration....	2	
Demonstrable plateau	2	
Total eliminated		32
Size of group for major study...		51

The twenty subjects eliminated from the major study on account of irregular attendance are not of interest in this connection. Three subjects were eliminated for extremely erratic performance in the tests. It is impossible that their real typewriting ability is even approximately represented by their erratic scores. The cause for their variability is undoubtedly due to lack of consistent interest in their work and in the tests. Most of the subjects took a competitive attitude toward the tests, the results of which were given them weekly by their instructor.

Ten subjects had learning curves which deviated from the typical hyperbolic form which we are here considering. This is a limitation of our method which is only applicable to the hyperbolic form of the speed-amount curve. Of these ten subjects five were eliminated from the major study on account of apparent linearity of the learning curves. *No learning curve can ever be continuously linear if it is plotted in the speed-amount form.* If it were linear the subject would have no physiological limit and he would in time reach the rather enviable attainment of infinite writing speed, which is of course absurd. Another alternative with a linear learning curve is that it is linear until it reaches the practice limit after which it remains at the limit. I can not entertain this as a possibility for it is inconceivable that an organic function like learning proceeds according to a linear relation until it bumps into some inflexible practice limit at which it stops and remains. The only possible explanation of apparently linear learning curves that I am willing to entertain is that they are in reality curved but that the degree of curvature is so small that it is concealed by the variability of the individual observations. Such learning curves are therefore indeterminate unless they be continued far enough to make the curvature appear in spite of the variations of the individual observations. This leads to the conclusion that the accuracy with which the learning coefficients can be determined is contingent on two principal factors. It varies a) with the degree of curvature of the learning curve, and b) inversely as the variability of the individual measurements. The coefficients of a learning curve with minimum variability may be determined with a minimum amount of visible curvature. The more variable the measurements the greater is the degree of curvature necessary for a fairly accurate determination of th learning curve coefficients. Whether the linearity of the curves of these five subjects is apparent or real can not be settled with the available data. If the linearity is real it constitutes a limitation in the use of the learning curve equation.

Two subjects were eliminated on account of delayed positive acceleration. Their learning curves constitute deviations from the usual shape of curve and can not be handled by the methods which we are discussing here. It is not certain that these measures are not simply cases of erratic performance.

Two out of the eighty-three subjects showed clear evidence of a plateau. Whether this is psychologically significant or simply due to the fact that these subjects were offered positions after attaining a specified typewriter proficiency is indeterminable. The higher order learning curve which followed the first curve is not carried far enough with either of these two subjects to justify determining the learning coefficients for the first and second order curves.

We have eliminated twelve out of sixty-three complete records. Generalizing from this fact we may conclude that the speed-amount form of learning for typewriting takes the hyperbolic form in about four cases out of five. This justifies our reference to it as the typical but not as the universal form of learning curve.

2) The Coordinates of Curves for Typewriting

My first intention was to plot the learning curves with speed as ordinates and time in weeks as abscissae, the speed-time form. Finding that the industry of the subjects during the practice hours varied immensely I decided that the psychological analysis would be more equitable if I measured practice in terms of total number of pages written rather than in terms of time, although time is statistically more readily obtained than the amount of practice. The practice sheets were all turned in to the teacher in charge who tabulated every week the number of pages written by each subject. According to the typewriter championship rules, attainment should be scored by deducing five words from the speed total for every error. For the purpose of psychological study I have separated errors from speed. The learning curves are all plotted as speed (words in four minutes, disregarding errors) against formal practice (pages written since entering the course). While errors are entirely disregarded in these curves the subjects were of course not informed on this point. The errors are studied separately by correlating them with the other learning characteristics. In this manner we shall arrive at a statement of the relationship between the several learning characteristics without artificially loading them with each other, as would be the case if we penalized the score for speed by the number of errors.

3) The Learning Coefficients for Typewriting

We shall use the following notation in studying typewriter learning.

X = *Practice,* in terms of the total number of pages written since entering the course.

Y = *Attainment,* in terms of the number of words written in four minutes.

x = Number of pages written when an individual test is taken.

y = Number of words written in a four minute test. It is the observed speed whereas Y is the speed indicated by the learning curve equation.

n = Number of tests taken.

y_a = Average speed in all tests or

$$y_a = \frac{\Sigma y}{n}$$

y_{20} = Average speed after twenty weeks of practice, and similar notation for the average speed at other stages of learning.

L = *Predicted practice limit* in terms of words written in four minutes.

R = *The rate of learning,* a constant which is numerically large for a low rate of approach and numerically low for a rapid rate of approach.

P = *Equivalent previous practice* in terms of practice units (pages written). It is the negative x-intercept of the speed-amount curve.

d = *Absolute deviation* which expresses the deviation of any single observation above or below the value indicated by the learning curve at that stage of learning. It is positive when the speed of any single test is above that indicated by the learning curve, and negative when the actual speed is below the learning curve. It can also be represented by the relation

$$d = y - Y$$

D = *Average deviation* for all tests during the year, or

$$D = \frac{\Sigma d}{n}$$

d_r = *Relative deviation* of an individual test, determined by the ratio

$$d_r = \frac{d}{Y} = \frac{y - Y}{Y}.$$

D_r = *Average relative deviation,* determined by the ratio

$$D_r = \frac{\Sigma(d_r)}{n}$$

V = *Coefficient of variability,* determined by the ratio

$$V = \frac{D}{y_a}$$

e = *Number of errors* made in an individual test as determined by the Typewriter Championship rules.

E = *Average number of errors* in all tests, or

$$E = \frac{\Sigma(e)}{n}$$

e_r = *Relative inaccuracy* of an individual test, or

$$e_r = \frac{e}{y}$$

E_r = *Average relative inaccuracy* for all tests, or

$$E_r = \frac{\Sigma(e_r)}{n}$$

A = *Coefficient of inaccuracy,* determined by the ratio

$$A = \frac{E}{y_a}$$

4) FINDINGS

a) *Writing Speed*

Figure 5 indicates the relation between average speed and number of pages practiced. The law of diminishing returns is shown by the continuity of the points representing average speed but we can not assert the universality of this form because we have already eliminated 12 out of the 63 available records. It will serve well as a norm of average performance for groups comparable with the one here represented. It is interesting to note that the curve does not pass through the origin. This is explainable by the fact that a person who has never written on a typewriter can, nevertheless, even on the first trial, make some finite score. I could not readily obtain a test from the fifty-one subjects prior to formal instruction on the typewriter. In order to ascertain what a truly initial score is, I asked ten of my students who had never touched a typewriter to take a four minute test. The mean as well as the average of these ten subjects was 27 words in four minutes. This is the point to which the composite learning curve in figure 5 projects.

The solid line of figure 5 represents the hyperbolic curve form. It is a good fit on the data which are represented by the small circles. The equation of the composite curve is

$$Y = \frac{216.(X+19.)}{X+148.} \qquad 40$$

in which Y is the average score for the 51 subjects and X is the number of pages of practice. The predicted limit of practice for speed is 216. words in four minutes, which agrees well with average typewriter speed. The average of the limits of practice as determined from the individual curves is 214. The equivalent previous practice for the composite curve is 19. This indicates that the general experience which these subjects brought to their first practice on the typewriter was equivalent to nineteen pages of formal practice. The rate of learning, R, for the composite curve is 129, a constant which varies inversely with the relative rapidity with which the limit of practice is approached. The average of the rates of approach as determined from the individual curves is 137. The composite curve, figure 13, is plotted against time in weeks instead of against amount of practice. It has an entirely different appearance. It shows the initial positive acceleration previously discussed.

We shall now turn to the individual records and ascertain by correlation methods the interrelations of the learning characteristics for typewriting.

The correlation between the predicted limit of practice and the average speed of writing for all tests during the year is 0.68. This correlation could not even theoretically be close to unity for while the fast writers tend to approach the higher practice limits there is considerable variation in the rate at which the limit is approached. It should be noted that the predicted limit was not attained by these subjects. The limit is predicted on the basis of the curve shape. If the predicted limit used here agrees at all closely with the ultimate writing speed of these subjects, the correlation of 0.68 between the practice limit and the average writing speed during eight months' instruction indicates that the latter measure is not a very reliable criterion of ultimate proficiency. I think that it constitutes another piece of evidence against hasty and self-confident predictions based on so called vocational mental tests.

The correlation between the predicted limit of practice, L, and the rate of learning, R, is 0.75. This indicates that those who approach a high limit of practice in speed generally approach their limit at a lower rate than those who are approaching a low speed as their limit. The regression is linear and hence the converse is true, namely that those who approach a low limit of practice generally approach their limit at a relatively higher rate than those who have a high limit of practice.

The correlation between predicted limit of writing speed and average number of errors in unit time is 0.10. Hence we conclude there is no discernible relation between the normal writing speed and the absolute number of errors made per unit time. But there is a noticeable relation between the coefficient of inaccuracy, A, and the practice limit, L. The coefficient of correlation is only —0.16 but the regression is non-linear. Those who approach a low practice limit for writing speed tend to be inaccurate, but those who approach a high practice limit are either accurate or inaccurate. On the other hand those who are unusually accurate tend to be fast writers whereas those who are inaccurate are either fast or slow writers. The number of subjects is not large enough to warrant the calculation of the eta coefficient.

The correlation between the predicted writing speed and the speed after eight weeks of practice is 0.27. Making use of the actual data instead of predicted performance, we find that the correlation between the speed after eight weeks practice and that after twenty weeks practice is 0.74. The correlation between predicted limit and the average writing speed for all tests is 0.68.

In this connection I wish to suggest what will perhaps be a more reliable technique in psychological prognosis. If a high degree of relationship can be established between the learning curve constants for a complex function, performance in which is to be predicted, and the corresponding constants for a simple learning function which can be completed during a single sitting, then the constants for the simple learning test would have diagnostic value in predicting performance in the complex function. Such coefficients would not be subject to the accidents of a first performance but would represent the organically more significant learning function as such. The diagnostic value of such a technique is largely dependent on the degree of difficulty of the material to be learned.

b) *The errors*

The relation between the average absolute number of errors made in each test and the number of weeks of practice is indicated in figure 16. This shows that the number of errors in unit time increases with practice. The relation between these two attributes may be expressed by the empirical equation

$$e = 0.12T + 2.1 \qquad 66$$

in which e is the average number of errors in a four minute test, and T is the number of weeks of instruction. The equation expresses a norm of average performance.

The analogous relation between average absolute number of errors and writing speed during the year is shown in figure 17.

This also indicates that the number of errors made in unit time increases with the attainment of writing speed. The relation may be expressed by the empirical equation

$$e = 0.023y + 1.5 \qquad 67$$

which is fairly representative within the limits of observation. However, the relative inaccuracy decreases with practice as indicated in figure 18, i.e., the number of errors per page decreases with practice but the number of errors per unit time increases with practice. There is no discernible relation between the relative inaccuracy (errors per unit time) and the rate of learning.

c) *The variability*

Those who have a high practice limit for writing speed usually have a larger average deviation from their learning curves than those who approach a low practice limit. This statement must be considered in connection with relation between predicted limit of writing speed and the average *relative* deviation. The correlation between predicted writing speed and average relative deviation for all tests is 0.27, indicating that the fast writers have a slight tendency to be more erratic in speed than the slow writers, even though the measure of variability is taken as the ratio of deviation to writing speed. According to this measure the writer of 60 words per minute is allowed a deviation from his representative learning curve twice that of a writer of 30 words per minute. But even according to this relative standard of variability the fast writers tend to be slightly more erratic in speed.

Figure 14 indicates that the deviations from the learning curve increase with practice but figure 15 shows that the ratio of deviation to theoretical writing speed, as indicated by the curve, decreases with practice. It is apparent that, just as one would expect, the absolute deviations increase with practice but the relative deviations decrease with practice. The decrease in the relative deviation with practice does not become noticeable until after about three months but after that the relation is approximately linear with practice time. In other words, the variability of the writing speed for any individual subject tends to decrease with practice, but if he is a fast writer he tends to be more variable in his writing speed than if he is a slow writer, even when the variability is measured in relative terms.

SUMMARY

1) Forms of the Learning Curve

We have discussed four different forms in which most learning data can be graphed. We have called these forms 1) the *speed-amount* curve, 2) the *speed-time* curve, 3) the *time-time* curve, and 4) the *time-amount* curve.

The *speed-amount* curve is plotted as speed, number of successful acts per unit time, or a multiple thereof, against the total number of formal practice acts, or some multiple of it. In plotting typewriter learning we have used words in four minutes, and the total number of pages written as the coordinates of the speed-amount curve. This form of curve is illustrated in figure 5 which gives average writing speed for fifty-one subjects against total number of pages written. The small circles indicate the observations and the solid line indicates the general trend of the learning. The solid line is represented by the general equation

$$Y = \frac{L(X+P)}{(X+P)+R} \qquad 20$$

in which

$L =$ Predicted practice limit in terms of speed units.

$X =$ Pages written.

$Y =$ Writing speed in terms of words in four minutes.

$P =$ Equivalent previous practice in terms of pages.

$R =$ Rate of learning, a constant which varies inversely as the relative rapidity with which the practice limit is being approached.

$K = P + K$, a constant used for convenience.

The particular line of figure 5 is represented by the equation

$$Y = \frac{216.(X+19.)}{X+148.}$$

which we may interpret as follows. The predicted average practice limit, L, for the group of fifty-one subjects is 216 words in four minutes or about 54 words a minute. The rate of learning, R, is a constant which in this curve has the value of 129.

Its only usefulness is in comparing the rates of several learning curves with each other. By itself, and for a single curve, it has no significance. When used to compare several learning curves the precaution must be observed that all curves so compared be plotted by the same units for the coordinates. The equivalent previous practice is nineteen pages. This is interpreted to mean that the general experience which these students brought to their first instruction on the typewriter was equivalent to writing nineteen pages on the machine. This coefficient as well as the predicted limit is based on the assumption that the unmeasured part of the learning before and after the observations followed the hyperbolic law. This assumption seems to be fairly safe since other learning curves, the actual observations for which start with practically zero attainment and continue almost to the practice limit, usually follow the hyperbolic form. See figure 11 which represents a substitution test learning curve carried almost to the practice limit, and the curve for subject No. 23 in figure 3 for typewriter learning which projects to the origin.

The *time-amount* curve is plotted as time, t, per unit amount of work against number of formal practice acts, X. Its equation is

$$t = \frac{X+K}{L(X+P)} \qquad 49$$

with notation similar to that of equation 20. Learning data can be changed from the time-amount form into the speed-amount form and vice versa by noting the fact that speed, Y, is proportional to the reciprocal of the time, t, per unit amount of work.

The *time-time* curve is plotted as time, t, per unit amount of work against total practice time, T. Its equation, 54, is derived from equation 20. This equation is too cumbersome for practical work and the speed-amount or time-amount curves should therefore be used unless one adopts a simple empirical equation for the time-time form.

The *speed-time* curve is plotted as speed, Y, against total practice time, T. Its equation, 55, is too unwieldy for practical work but it serves to demonstrate the following proposition regarding positive acceleration. If we assume that the typical form of speed-amount curve is hyperbolic, then the learning curve in the speed-time form must necessarily have an initial positive acceleration which changes to a negative acceleration when the attainment has reached one-third of the limit of practice. The positive acceleration can not, of course, be obtained when the initial score is greater than one-third of the practice limit. It can only be observed when the initial score is less than one-third of the practice limit.

The influence of different values for the learning coefficients on the shape of the learning curve may be summarized in the following comparisons. Figure 7 shows two learning curves approaching two different limits at the same rate. Figure 8 shows two curves approaching the same limit at different rates. Figure 9 shows two curves approaching the same limit at the same rate but differing in the amount of previous practice. Curve A has a start of forty practice units over curve B. Figure 10 shows two curves, one approaching a high limit at a low rate, the other approaching a lower limit at a high rate. The important feature of this comparison is that one who learns rapidly but with a low limit will do better in the first stages of the learning than one who learns slowly with a high limit. The comparison shifts later in favor of the learner with the high limit. This is a condition which experimenters on learning should be on the look-out for in order to guard against the erroneous comparison of two subjects from insufficient practice data.

2) Outline for Calculating the Learning Coefficients

We shall describe two methods of calculating the coefficients. These are 1) the method of all points, and 2) the method of three points.

1) *Method of all points:*

1) Arrange the data in two columns as follows: X, the total number of formal practice acts since the beginning of practice, or a multiple of this number, and Y, the number of successful acts in unit time, or a multiple of this number. The multiple

used for the Y-column need not of course be the same as that for the X-column.

2) Draw a chart analogous to figure 5. Leave room for a negative x-intercept. Always include the zero point of the y-scale on the chart.

3) Select a representative initial score. In figure 5 the actually observed initial score, 27, was used. Denote this by the symbol Y_1.

4) Compute the values of $X/(Y—Y_1)$ and $XY/(Y—Y_1)$ for each observation. Arrange these in two columns.

5) Draw a chart analogous to figure 6 with the coordinates determined in step 4. If the data so plotted fall nearly on a straight line the speed-amount curve is hyperbolic. If it does not, the use of the learning curve equation is not justified and other methods must be resorted to.

6) Fit a straight line through these points in figure 6 in such a manner that there are about as many points on one side of the line as there are points on the other side. This procedure is called "rectifying the equation." The line can be fitted more accurately by the method of least squares but since that is a rather laborious procedure it should be avoided unless the points are so badly scattered that they can not readily be fitted by inspection. Even then it is doubtful whether one is justified in applying the equation to learning data so erratic that a straight line can not be fitted by inspection.

7) Continue this line until it intersects the x-axis. The x-intercept gives the value of K/Y_1, and since the value of Y_1 is already known the value of K can be readily determined.

8) The slope of the line is numerically equal to the predicted limit, L.

9) The constant P may then be determined from the equation

$$P = \frac{K \cdot Y_1}{L}$$

10) The constant R is then determined by the equation

$$R = K - P$$

since K and P are known.

2) *Method of three equidistant points.*

The first two steps of this method are identical with the first two steps in the method of all points.

3) Draw a *smooth* curve through the observations. If the data show irregularities in the rate of learning draw the smooth curve so that it has approximately as many observations above the line as there are observations below the line. A ragged line through all the more or less erratic observations will not serve the purpose. If the smooth curve representative of the data is not of the hyperbolic form the method of this learning curve equation is not applicable.

4) Select the three following points:

$X_1; Y_1$

in which X_1 is zero, and Y_1 is the ordinate to the smooth curve at this value of X. If the learning data have no irregularities the value of Y_1 will be identical with the initial score.

$X_2; Y_2$

in which X_2 is one half of the total amount of formal practice and Y_2 is the representative ordinate to the curve for this value of X.

$X_3; Y_3$

in which X_3 is the total amount of formal practice and Y_3 is the ordinate to the smooth curve at this value of X. If the learning data show no irregularities, this will be identical with the final score.

5) Determine the numerical value of the constant K from the equation:

$$K = \frac{X_3(Y_2 - Y_3)}{Y_3 + Y_1 - 2Y_2}$$

6) Determine the numerical value of constant L by means of the following equation:

$$L = \frac{Y_3(X_3 + K) - Y_1 \cdot K}{X_3}$$

7) Determine the numerical value of the constant P by the following equation:

$$P = \frac{Y_1 \cdot K}{L}$$

8) Determine the numerical value of the constant R by the following equation:

$$R = K - P$$

After the constants of the learning curve equation have been numerically evaluated it is best to check the arithmetical work by computing the theoretical value of the attainment for one or two points according to the following formula. These theoretical values of attainment should not differ much from the actually observed values unless the learning has been very erratic.

$$Y = \frac{L(X+P)}{(X+P)+R}$$

A gross measure of the variability of the learning may be determined from the equation

$$V = \frac{D}{y_a}$$

in which D is the average deviation from the theoretical curve for all the observations, and y_a is the average attainment as determined from all observations.

3) Typewriter Learning

The following relations were found to be significant with regard to typewriter learning.

The correlation between the predicted practice limit and the average writing speed for all tests which covered about seven months is +0.68. The correlation between practice limit and rate of learning is +0.75. This indicates that those who approach a high practice limit usually do so at a lower rate than those who approach a low limit since the coefficient, R, for the rate of learning varies inversely with the rate of learning. There is a noticeable relation between accuracy and the predicted practice limit. Those who approach a low practice limit for writing speed are usually inaccurate, but those who approach a high practice limit are either accurate or inaccurate. On the other hand those who are inaccurate are either fast or slow writers. The number of subjects, 51, is not large enough to warrant the calculation of the eta-coefficient. The correlation between the predicted practice limit and speed as determined in the test at the 8th week is $+0.27$. There seems to be no relation between the predicted writing speed at the limit of practice and the number of errors made in unit time.

The number of errors made in unit writing time increases with practice. Similarly there is a positive relation between the number of errors in unit time and writing speed during practice. However, the number of errors per unit amount of work decreases with practice. See figures 16, 17, and 18. There is no discernible relation between the relative accuracy and the rate of learning.

Those who have a high practice limit for writing speed usually have larger relative deviations from their theoretical learning curves. The variability of the writing speed for any individual subject tends to decrease with practice, but if the student is a fast writer he tends to be more variable in writing speed, than if he is a slow writer, even when the variability is measured in relative terms. According to this standard of variability the writer of 60 words per minute is allowed a deviation from his representative learning curve twice that of a writer of 30 words per minute. But even according to this relative standard of variability the fast writers tend to be slightly more erratic in speed.

Considerable ambiguity in discussions about learning curves has been caused by the comparison of learning curves with different units for the coordinates. Thus we are entirely safe in saying that the speed-amount curve is never continuously linear. It would lead to infinite speed of performance which is of course absurd. But while that statement is obviously true it does not entitle us to jump to the denial of say linear error-time curves. It is quite possible for errors plotted against time to be linear. Therefore we should always specify the coordinates for the curves we are discussing.

While I have confined myself throughout to what I have called the typical hyperbolic form of the speed-amount curve it is quite essential to keep in mind that this form of curve is not universal and that consequently it is impossible to make sweeping generalizations except in so far as we explicitly limit ourselves to the relations which follow from the assumed hyperbolic form with which we started.

The preceding pages have been filled with so much algebraic manipulation that the reader who has long since dropped the algebraic thinking of his school days may find their very appearance formidable and distasteful. For the benefit of those who have acquired an aversion against symbolic notation I wish to call attention to the outline for calculating the coefficients and the section on learning curve forms in the summary. In those sections will be found all that is really essential in applying the method. If the use of empirical equations in the quantitative study of the multifarious aspects of memory is at all furthered by the present study I shall be content though the particular forms used here are superseded by others.

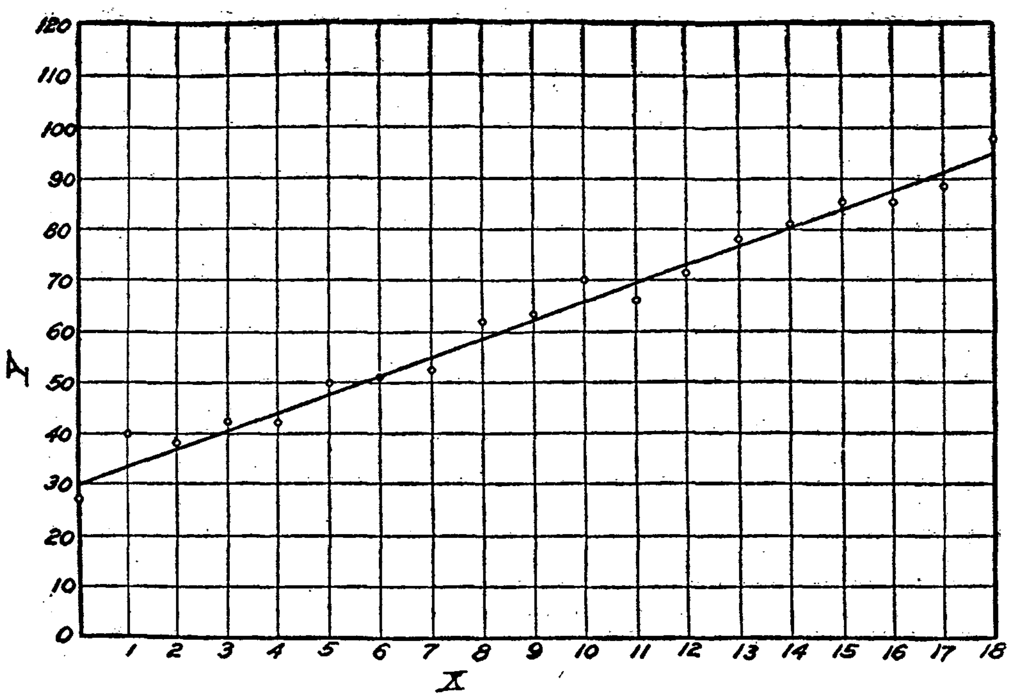

FIGURE 1

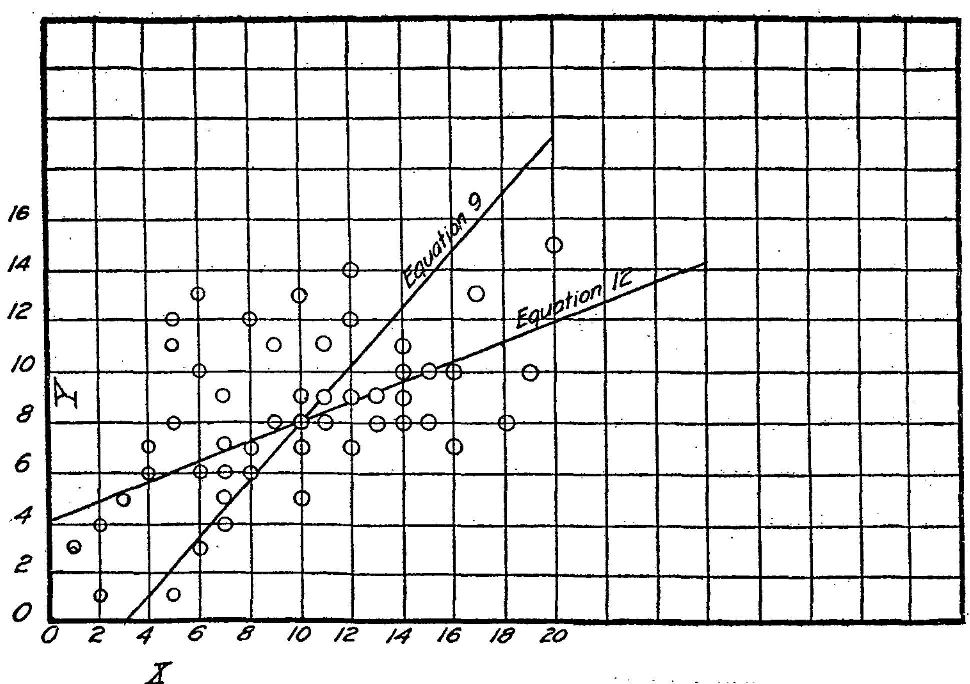

FIGURE 2

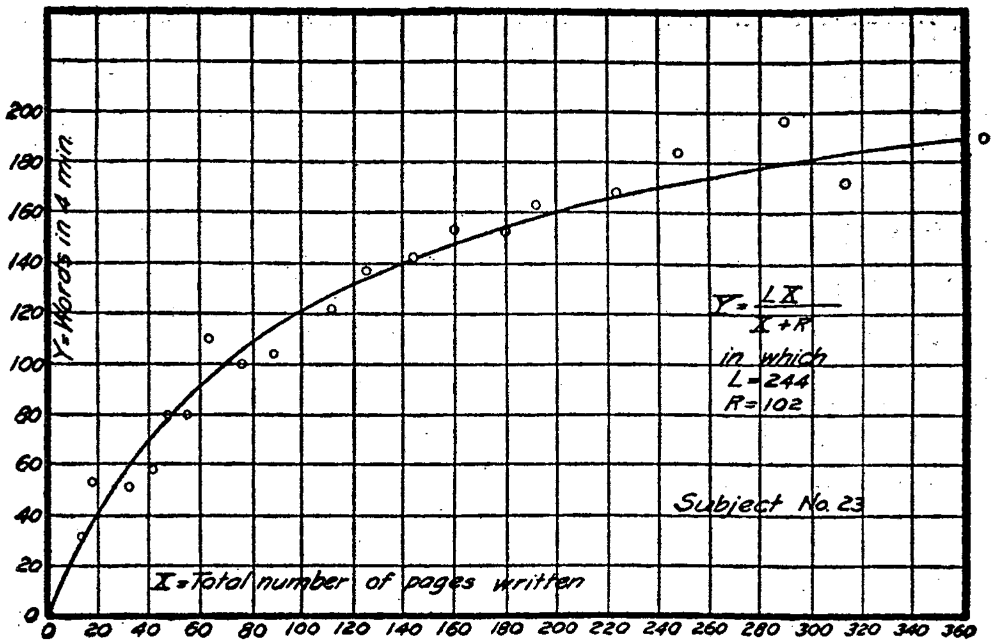

FIGURE 3

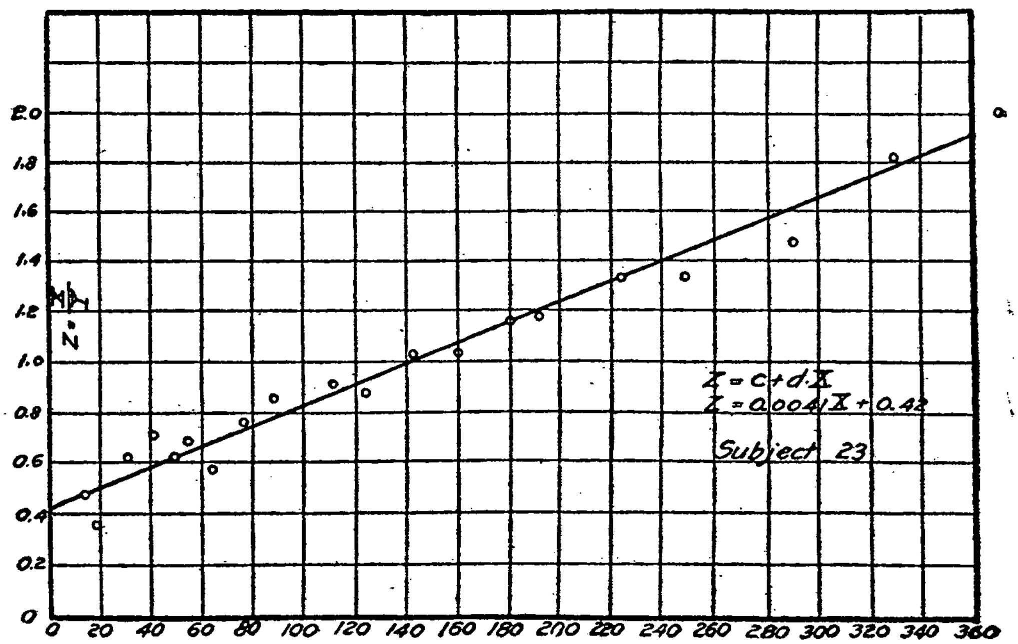

FIGURE 4

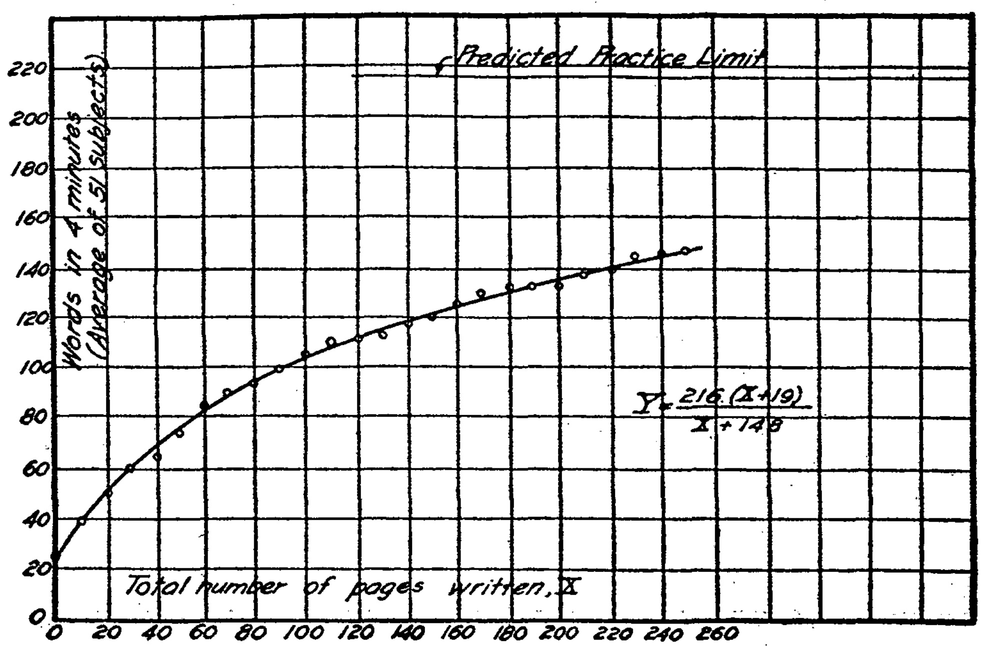

FIGURE 5

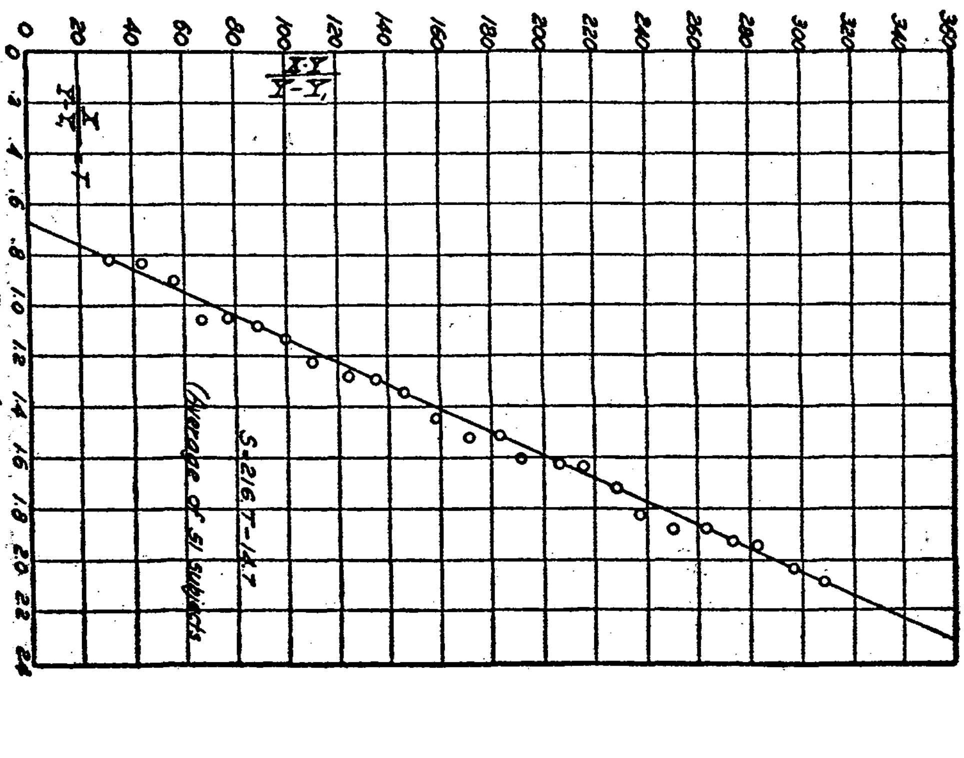

FIGURE 6

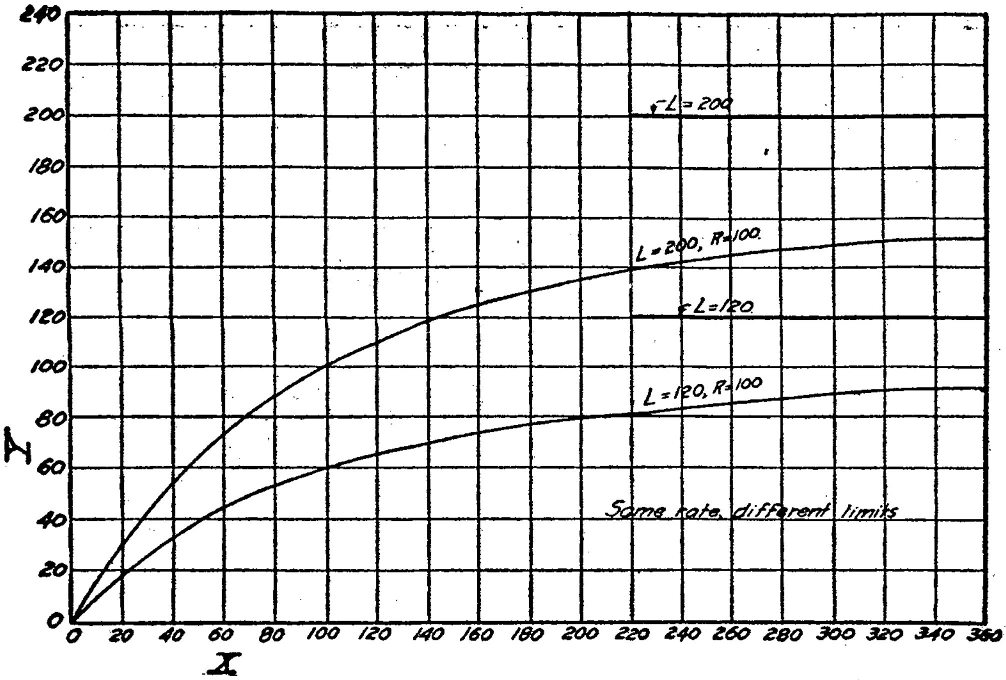

FIGURE 7

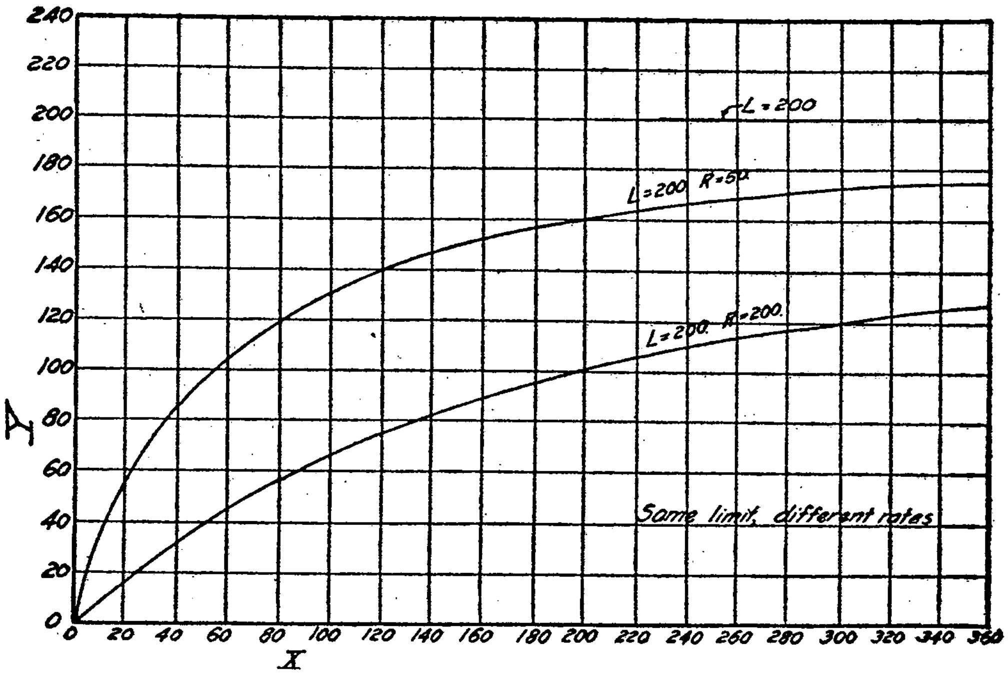

FIGURE 8

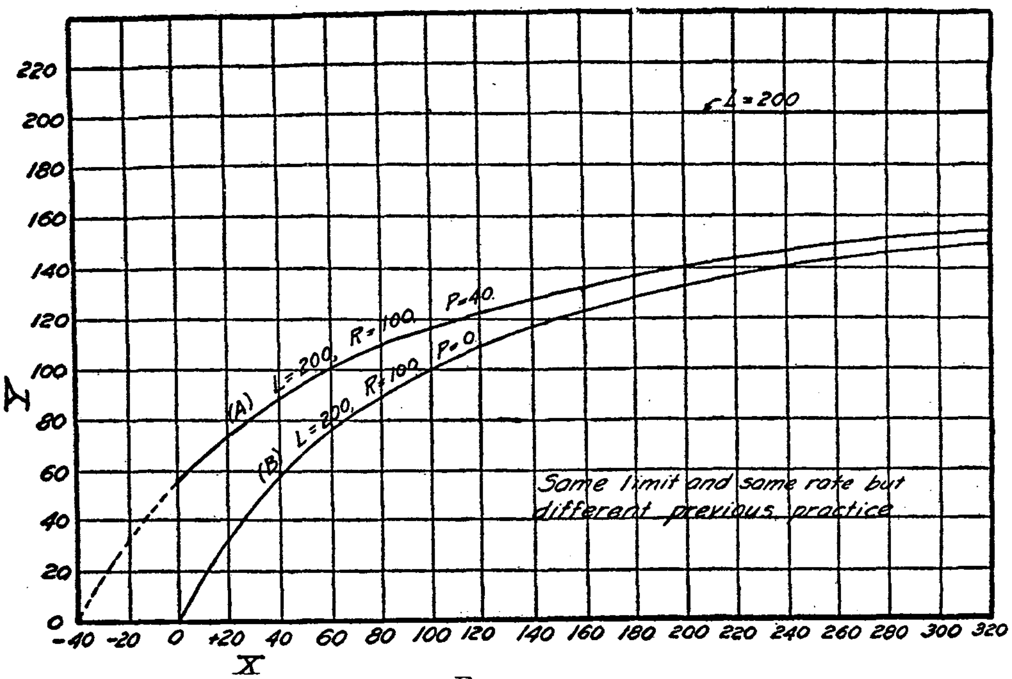

FIGURE 9

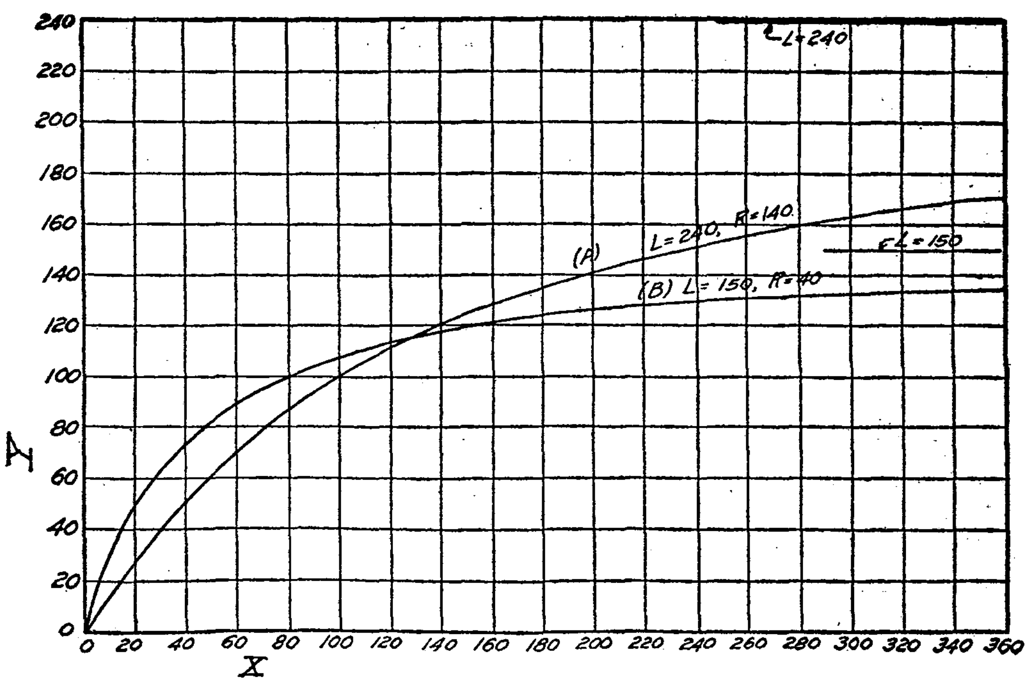

FIGURE 10

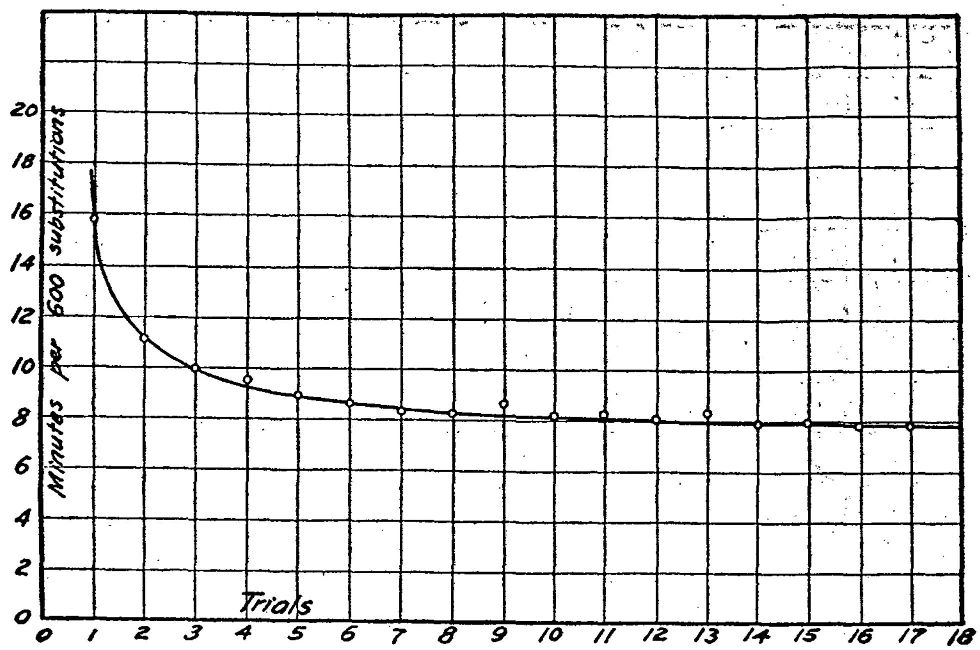

FIGURE 11

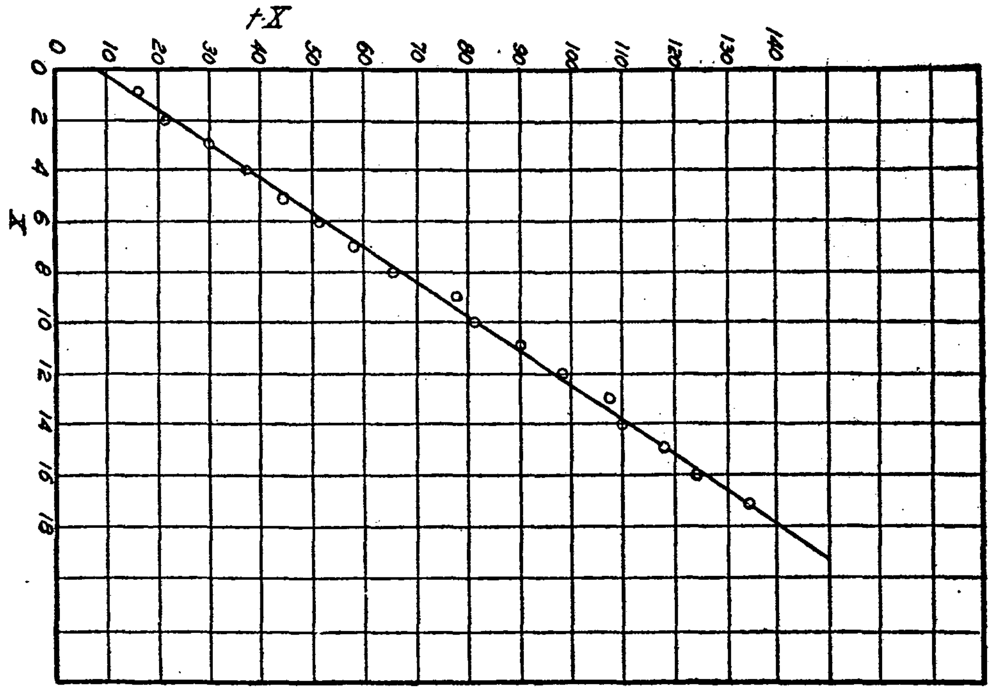

FIGURE 12

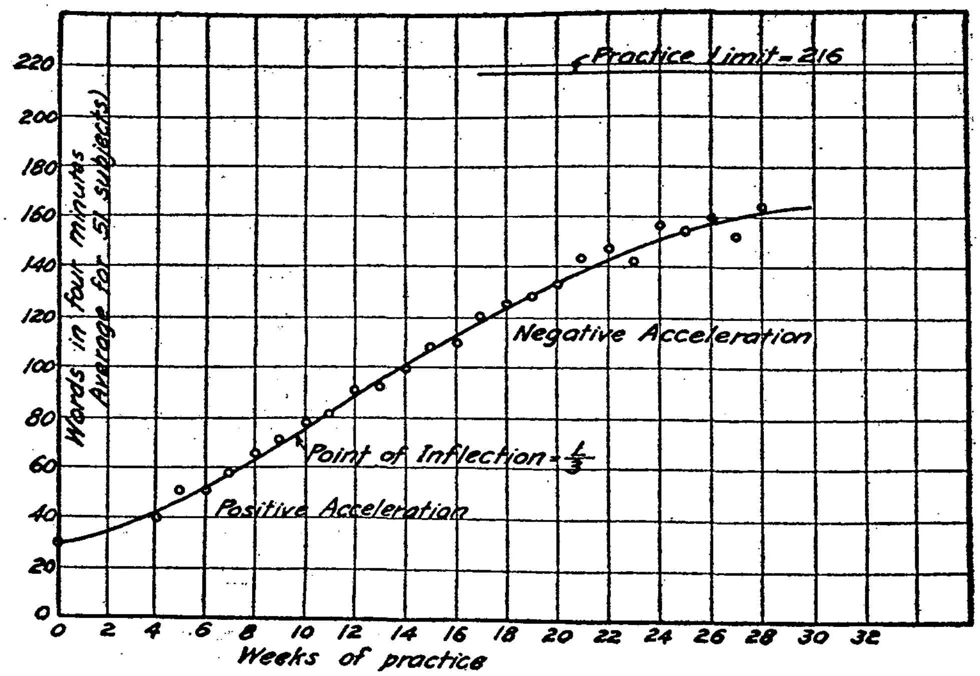

FIGURE 13

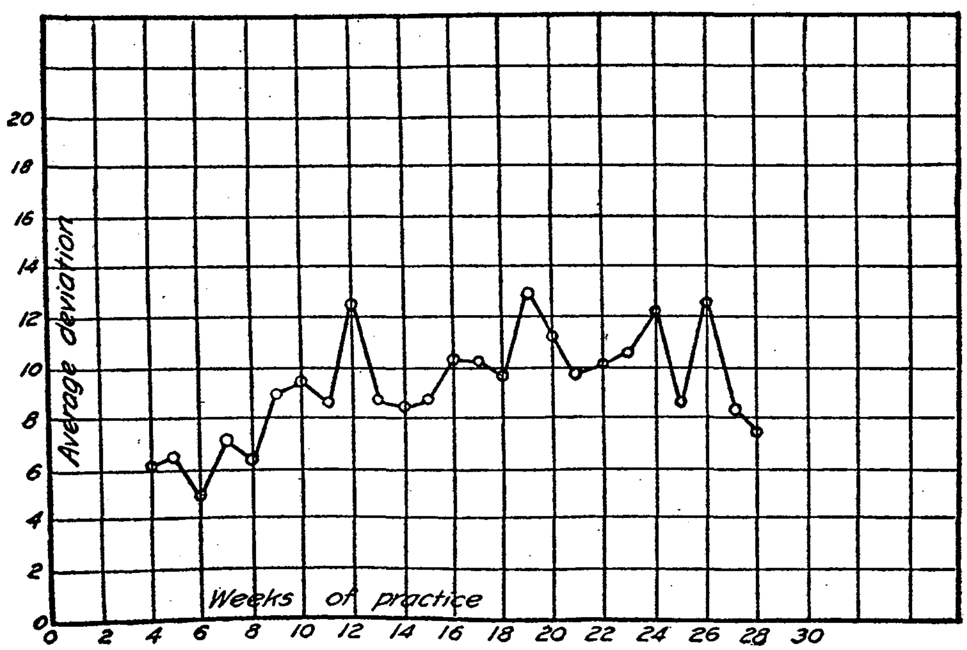

FIGURE 14

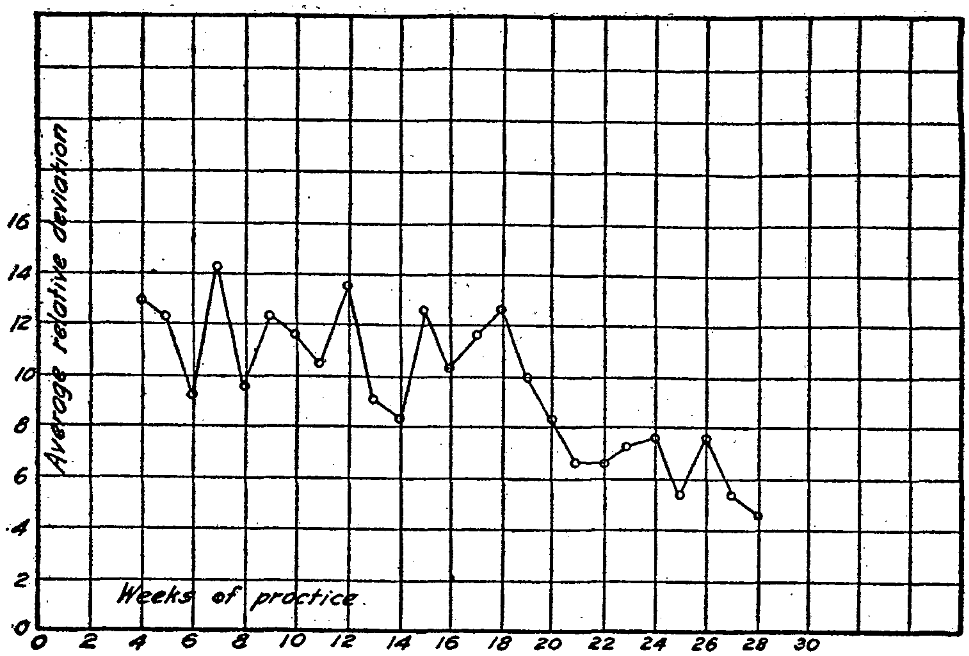

FIGURE 15

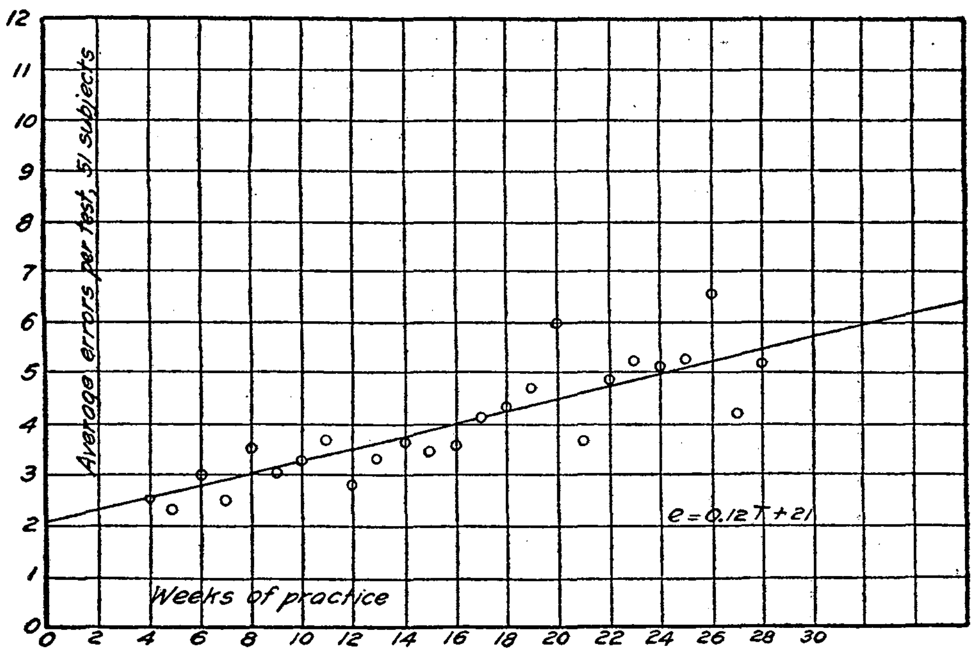

FIGURE 16

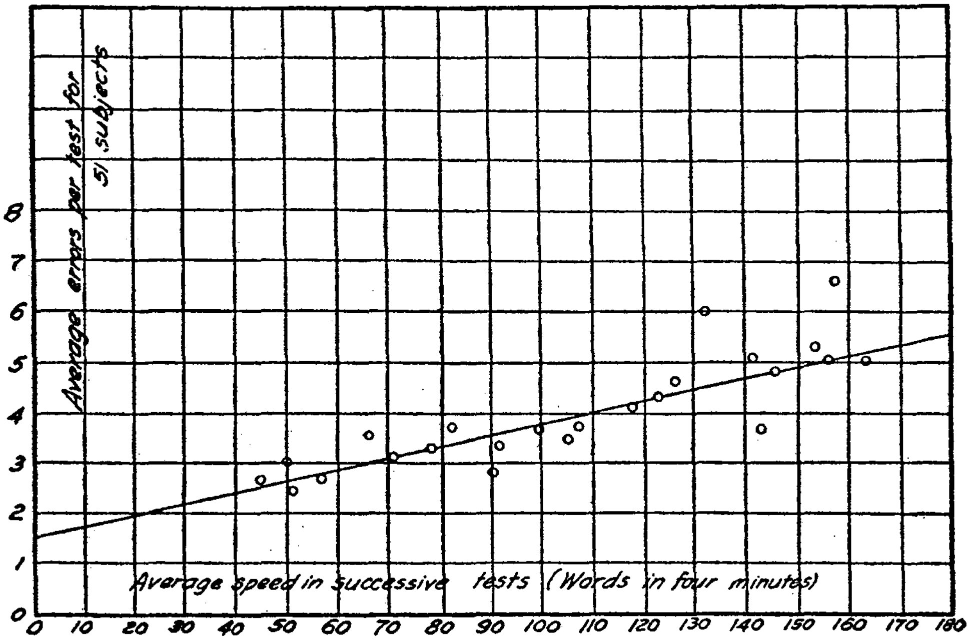

FIGURE 17

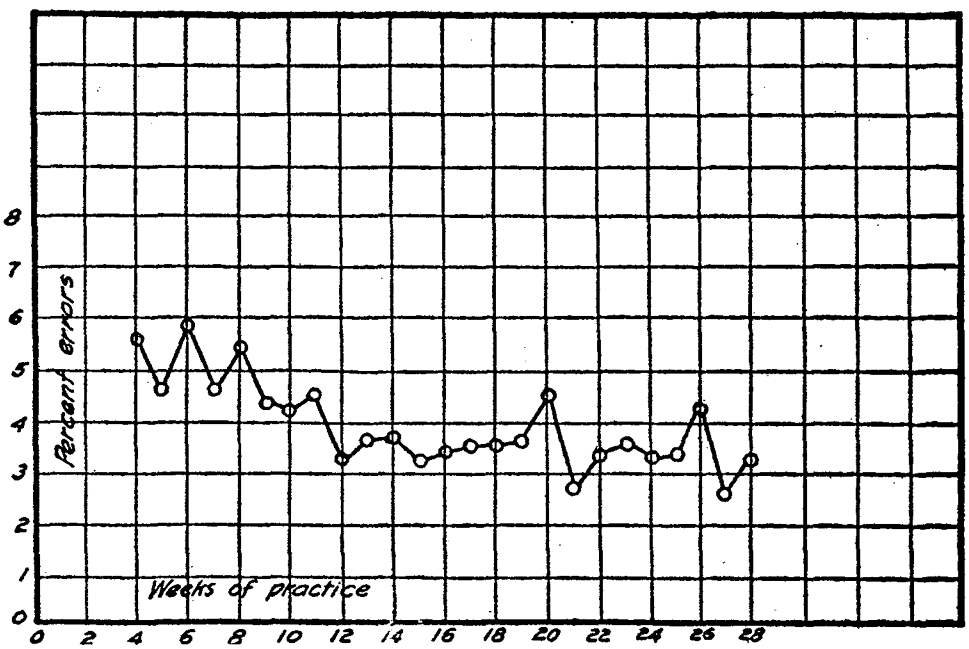

FIGURE 18