

AD-A034 041

HUMAN RESOURCES RESEARCH ORGANIZATION ALEXANDRIA VA  
ACQUISITION OF A MULTIDIMENSIONAL RESPONSE (Y), WHERE DURATION --ETC(U)  
APR 60 J F FOLLETTIE

F/G 5/9

DA-49-106-QM-1

NL

UNCLASSIFIED

| OF |  
AD  
A034041

NOTE  
THIS  
FILM

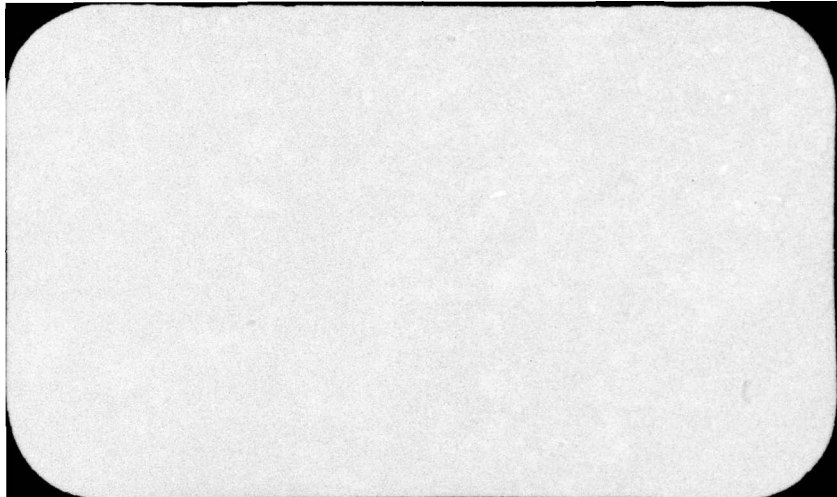


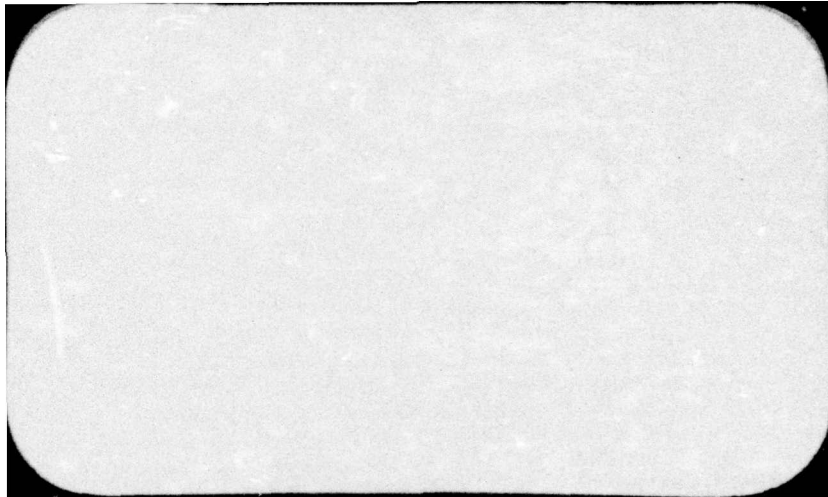
END

DATE  
FILMED  
2 - 77



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A





RESEARCH MEMORANDUM

ACQUISITION OF A MULTIDIMENSIONAL RESPONSE (Y), WHERE  
DURATION OF ACQUISITION (X) AND ASSESSMENT (t)  
INTERVALS ARE NON-FIXED

By

Joseph F. Follettie

April 1960

Approved:

*Preston S. Abbott*

PRESTON S. ABBOTT  
Director of Research

U. S. Army Infantry  
Human Research Unit  
Fort Benning, Georgia

Number 20

This document has been approved  
for public release and sale; its  
distribution is unlimited.

ACCESSION for	
RTIS	White Section <input checked="" type="checkbox"/>
DOC	Buff Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION.....	
.....	
BY.....	
DISTRIBUTION/AVAILABILITY CODES	
Dist.	AVAIL. and/or SPECIAL
<b>A</b>	

Copy \_\_\_\_\_

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) <span style="border: 1px solid black; padding: 2px;">(6) Acquisition of a Multidimensional Response (Y), Where Duration of Acquisition (X) and Assment(t) Intervals are Non-Fixed.</span>		5. TYPE OF REPORT & PERIOD COVERED <span style="border: 1px solid black; padding: 2px;">(9) Research Memorandum</span>
7. AUTHOR(s) <span style="border: 1px solid black; padding: 2px;">(10) Joseph F./Follettie</span>		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Human Resources Research Organization (HumRRO) 300 North Washington Street Alexandria, Virginia 22314		8. CONTRACT OR GRANT NUMBER(s) <span style="border: 1px solid black; padding: 2px;">(45) DA-49-106-qm-1</span>
11. CONTROLLING OFFICE NAME AND ADDRESS Office of the Chief of Research & Development Department of the Army Washington, D.C. 20310		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS <span style="border: 1px solid black; padding: 2px;">(12) (45) P.</span>
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE <span style="border: 1px solid black; padding: 2px;">(11) Apr 1969</span>
		13. NUMBER OF PAGES 40
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES Research performed by HumRRO U. S. Army Infantry HRU (now HumRRO/Columbus) under Project BASICTRAIN.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Performance Multidimensional Response		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) In this report, a multidimensional view of performance is discussed, conditions fair to expression of student rate differences during acquisition and testing are elaborated and a model which is consonant with multidimensional performance and exploitation of differences in student rate preference is presented.		

## PREFACE

When, in the course of inquiry, and to what extent should the investigator engage in formal model-building in order to influence the course of research? Undoubtedly, the answer will vary with complexity of the phenomenon under study, research auspices, and personal taste. In at least one sense, it would seem that the modelling requirement becomes more pressing as one moves nearer and nearer to a commitment to solve problems actually to be found in the everyday world. The everyday world supplies performance requirements which are not just academic. I will argue that these requirements usually define performances which are, in some sense, multidimensional, necessitating what I will call response modelling.

The model to be presented--which stems from conceptual requirements of Task BASICTRAIN--grew out of a minuscule problem in research on transmission of essentially rote learning contents. Given a group of students, each utilizing as much exposure time (X-axis value) as he desired and each yielding a performance (Y-axis value) based on his own unique exposure time, how could one summarize the data? Particularly, how could one summarize the data if his preference ran to a curvilinear view of the form of the function Y of X instead of the unbelievable linear view so often adopted in treating the data?

Once embarked upon solving this problem, it became evident that a little extra effort would yield a model which better represented

performance requirements underlying the research at hand. Given such a model, determination of the empirical requirement and consequent setting of the experimental paradigm or paradigms is readily accomplished, and this I have done.

One of the great obstacles to the conduct of systematic investigations into human learning--specification of the units of content and associated input--is not touched here. Until human learning research can, appreciably, be emancipated from response-definition and intuitive scaling of independent variables and from the practice of describing only an infinitesimal fraction of the stimulus complex underlying investigations, all efforts at quantification of learning will be as mighty edifices built upon foundations of jelly. If it is too much to hope that a neat centimeter-gram-second system of units will be found appropriate to inputs to which we are irrevocably committed, at least we must push beyond the arid regions of attribution, adjectivity, and contingency when scaling content and associated input. Although this paper is generally mute with regard to our Z-axis units, crude beginnings on a more satisfactory dimensionalization of the stimulus complex are underway within the BASICTRAIN framework.

## CONTENTS

	Page
PREFACE-----	11
I. INTRODUCTION-----	1
A. Performance: One Dimension or More?-----	1
B. The Fair Conditions to Aptitude Maximization in Education-----	5
II. DIMENSIONS OF RESPONSE Y-----	7
A. The System in Which Y Is a Consequence-----	7
B. Our View of Y-----	9
III. RELATIONS BETWEEN $X$ , $s$ , $s/t$ , AND $D$ : A PRELIMINARY • MODEL FOR STUDIES IN EDUCATION-----	13
A. Why Model Now?-----	13
B. Terms and Relations-----	14
C. Evaluation of Empirical Constants $k$ , $C$ , $a$ , $b$ , $m$ -----	18
D. Steps in Generation of $Y = [1 - e^{-b(C+T)}] / mX$ -----	22
E. Retention Over Time ( $D \neq 0$ )-----	30
IV. SUMMARY-----	34
REFERENCES-----	35
AFTERTHOUGHT-----	36

ACQUISITION OF A MULTIDIMENSIONAL RESPONSE (Y), WHERE  
DURATION OF ACQUISITION (X) AND ASSESSMENT (t)  
INTERVALS ARE NON-FIXED

I. INTRODUCTION

In studies of human behavior, two tendencies stand out. First, performance is widely conceived and measured as if it were unidimensional. Second, the tendencies of students to diverge in favored content presentation and acquisition demonstration rates are widely ignored or stifled.

In this paper, a multidimensional view of performance will be discussed, conditions fair to expression of student rate differences during acquisition and testing will be elaborated, and a model which is consonant with multidimensional performance and exploitation of differences in student rate preference will be presented.

A. Performance: One Dimension or More?

Before one can address the question of dimensions of response R, the question of the domain of R must be treated. It is conceivable that the number and kind of dimensions of R which will be required for an adequate description will vary with the domain of R--its behavioral scope. A spatially and temporally huge R--say a unitary response sequence which takes five minutes, involves radical shifts in respondent topography, and moves the respondent a distance of 100 yards from origin--will generate a more elegant descriptive scaffolding than a much more modest--molecular, if you will--response. Behavior scientists by no means agree that freedom to set the scope of performance of interest is unrestrained.

Educators address--among other objectives--a requirement that behaviors which are taught impact upon the halls of commerce. The question of validity of response specification has empirical urgency when it is asked with respect to an everyday world in which decisions must be made and alternate actions are not equally consequential. The unhurried investigator in the haunts of untrammled science need not accept a requirement that his impact upon commerce be immediate or direct. Hence, validity of response specification in untrammled science need not be established with reference to the world of commerce.

Brunswik (1) laid down the requirement that laboratory conceptualizations of behavior and its conditions be representative of life situations to which laboratory findings will be directed. To follow the lead of Brunswik is to define laboratory tasks in terms of everyday world considerations. I have no position on how generally behavior scientists should bow to Brunswik's dictum on representative design. Those of us who, by virtue of interest or charge, are rather immediately concerned with the question of efficient education for worldly action will find it difficult, however, to exclude a representativeness plank from our platform. Given the dictum of representativeness-findings will be applicable to behaviors in the worldly domain--we will proceed to specification of a worldly R and its laboratory representation. It will be my position in this paper that most utilitarian R's of everyday life cannot be characterized adequately unless the characterization is multidimensional.

In some sense, most of us could agree that all learning situations compel or at least invite the measurement or assessment of more than one effect of acquisition. If many investigators and theorists treat performance as unidimensional--among educational investigators the test score scale is by far the most popular unidimensional preference--it hardly taxes the mind to supply alternative or supplementary indicators of acquisition.

In addition to the score, or accuracy, dimension--itself potentially multidimensional--the definer of response has access to several varieties of response rate, to response probability, and to response attenuation over time. Skinner (6) defined the response of operant conditioning in terms of rate of responding and number of responses to extinction. Hull (4) presented a somewhat more elaborate view of response. In this particular system of Hull--his systems vary--there are four indicators of an over-all reaction to R. These are:

$p$  = the probability of R

$s_r^t$  = the latency of R

$n$  = the number of unreinforced R's to extinction

$A$  = amplitude of R

Each of these indicators is, for Hull, a specified function (Postulates 12-15, p. 344) of theoretical constructs which in turn are specified functions of input. As Koch notes--in Estes et al. (2)--Hull fails to specify topography of the data-level correlate of R.

R has four properties, but what is R? Worse yet, from a standpoint of the representative designer, Hull fails to convince us that the four indicators represent a systematic view of the requirements for adequate description of the system concept of R. Even so, we see in Hull's system a program of response definition in keeping with representative design notions of performance complexity. In it, R has become something more than whatever happened to be at hand.

Many investigators at least implicitly acknowledge an interest in a multidimensional R by virtue of the fact that results are reported along two or more dimensions of behavioral measurement. Most studies dealing in multiple response measures can best be thought of as directed at an eventual multidimensional definition of R. We are accustomed to being told that if R is taken as R', then R is Function A of X, while if R is taken as R'', then R is Function B of X. While such an approach does not get the job done as I would have it done, it hints, as an iceberg, of more substance than meets the eye.

It would seem useful not to consider R as multidimensionally defined until its author is willing to assert two things: (1) the dimensions of R, in kind and number, which underlie adequate description of R (e.g., R', R''), and (2) the functional relation between the dimensions of R (e.g.,  $R = R' \cdot R''$ ). Of course, topographic delineation of R is propaedeutic to useful dimensionalization of R. (First take its picture and only then analyze it.) Delineation of response topography will be treated only implicitly in this paper.

B. The Fair Conditions to Aptitude Maximization in Education

Recently, the conductors of behavioral research for education have discovered the rate dimension of the genus Aptitude. This discovery has given way to a spate of studies wherein the student is removed, during acquisition, from the grip of a classroom rate of content presentation not his own. Some may feel that with these studies the human student employed in behavioral research for education has been extended all the privileges allowed to rats, mice, and monkeys participating in behavioral investigations. It is my position that this conclusion is incorrect. His neighbors-down-the-tree still enjoy one advantage not generally--if at all--accorded our student. Not only does the rat, the cat, or the canary most usually respond at his own rate to a stimulus complex during acquisition, thus controlling rate of presentation, but he also responds at his own rate when it comes time to demonstrate the extent of his acquisition. That is, he controls training time (X) and testing time (t). The human student employed in behavioral research for education has at last been allowed to fix X per unit presentation. Not so, t.

It seems reasonable that an individual whose aptitude for a task is low will be slow to acquire and slow to demonstrate acquisition. There is little point in allowing a student to spend a good deal of time in training and then to cut him off during testing before he has had an opportunity to demonstrate acquisition to the full extent of the requirement posed by the test instrument. Particularly when we

are concerned with demonstrating efficacy of exploitation of some individual difference in aptitude, we cannot hope the test will be fair if the student is considered to possess his rate of response during acquisition and my rate during testing. For the student to control X and the experimenter to control t is to employ an aptitude test to measure achievement.

Absence from the learning literature of studies which feature simultaneous student control over X and t may arise from an undue statistical orientation on the part of investigators. Speaking of and to the conductors of behavioral research for education only, we may be tailoring the experimental learning situation and its response requirement not to the world we live in but to the world of available statistical and research design tools. This is unpardonable technocracy. It is fitting the nut to the wrench instead of the wrench to the nut.

## II. DIMENSIONS OF RESPONSE Y

Thus far, we have established that a response R--a behavior of ultimate interest--will be utilitarian activity in a reference world of commerce. We contrast a utilitarian R with a creative one of the sort characteristic of the domain of the plastic arts. Moreover, we have committed ourselves to the notion of a response Y--the analogue of response R--whose domain will be the classroom and the laboratory. While we cannot settle here the conditions under which findings generalize or, in fact, what is meant by generality, we have in principle accepted the Brunswikian principle of representative design. Hence, response Y and its conditions will be taken as representative of response R and its conditions. Finally, we have asserted that R--and hence Y--quite probably cannot be described adequately unless more than just one of its properties enters into the description. Our next task is to describe the conditions which give Y meaning, to give some idea of the scope of Y, and to specify its apparent necessary and sufficient dimensions for purposes of adequate description.

### A. The System in Which Y Is a Consequence

In any acquisition-demonstration situation, we may distinguish two general classes of response. There are responses associated with acquisition itself, which culminate in student exposure to a content of interest, and there are responses associated with demonstration of acquisition. Response Y will belong to the demonstrational class. As such, it will occur during testing time.

The test in which Y will occur will consist of discrete items of written or voiced material which, taken together, constitute an acceptable sample of the content presented during an acquisition phase. Its items will be generally equally difficult, reflecting a content whose items of information possess the same characteristic, for how else can equally difficult test items be equally valid?<sup>1, 2</sup>

Temporally, the acquisition-demonstration situation surrounding Y will consist of the following:

X = a training interval of student-determined duration, which coincides with one negotiation of a Content A, the content to be learned

D = a post-training interval of investigator-determined duration, which coincides with one negotiation of a Content I, a content which interferes with Content A

---

<sup>1</sup> The assumption to which one is entitled regarding item difficulty is a function of the nature of the content taught. It is not properly a function of how difficult it is to devise items of a specified relative difficulty, although the psychometrically-oriented often talk as if it is. One hears many prohibitions against the notion of test items of equal difficulty, on grounds that it is almost impossible to construct such items. Perhaps there are few occasions which warrant test items of equal difficulty. The reason for this is not construction difficulty but that the curriculum does not warrant the approach.

<sup>2</sup> The notion of item difficulty as I employ it refers strictly to item characteristics. While it might seem quite tautological to relate that I mean by item difficulty item difficulty, it needs saying. In the hands of the psychometrically-oriented--for instance, Guilford (3)--the concept becomes a joint function of item characteristics and student past experience. Eventually, I plan to elucidate a procedure--consistent with the model to be presented--which removes past experience from establishment of item difficulty.

t = a testing interval of student-determined duration, which coincides with one negotiation of a Content A', an acceptable sample of Content A, given in test format.

Certain treatments will be associated with the content to each class of interval, namely:

ZnA = acquisition treatments of Content A (e.g., the Z-axis dimension redundancy, n values 1, 2, 4, etc., applied to A)

ZnI = interference treatments of Content I

ZnA' = testing treatments of Content A' (e.g., the Z-axis dimension response configuration, n values recognition, recall, relearning, etc.), where the relation of A' to A in kind is understood and in extent is specified

The symbol s will stand for test score, defined as follows:

$$\underline{s} = \frac{\text{number of items correct}}{\text{number of items in test}}$$

The outstanding feature of the system in which Y is a consequence is that it features student-to-student variation in X, t, and s.

#### B. Our View of Y

To recapitulate in terms of a three-dimensional coordinate system, where D = 0 (offset of X followed immediately by onset of t, absence of Content I), we have:

On the Z-axis, the independent variable (say redundancy).

On the X-axis, duration of exposure to Content A under conditions of the independent variable; that is, the interval X.

On the Y-axis, duration of exposure to Content A' and score on Content A': the interval t and score s.

It should be noted that the Y value which occurs under the condition  $\underline{D} = 0$  is properly called  $\underline{Y}_{\underline{D} = 0}$ .

The world of commerce and, hence, the world of education generally require occurrence of a Y of pith and moment under the condition  $\underline{D} = n$ , where the interval D is cluttered with all sorts of interfering Contents I. Hence, a representative design will incorporate such features as are to be found in response attenuation (forgetting) designs.<sup>1</sup> Later, we will look at Y under conditions of  $\underline{D} = 0$  to  $n$ . For the present, Y will be defined under the more artificial condition of  $\underline{D} = 0$ .

For a given value of the independent variable Z,  $\underline{Y} = f(\underline{s}, \underline{t}) = f(\underline{X})$ . Let us translate this as "The measure of performance equals a function of a measure of accuracy and a measure of speed which in turn are functions of duration of exposure to the acquisition treatment."

Accuracy of performance is reflected by the dimension of s--here an equal interval scale ranging from 0 to unity. The minimum requirement for defining accuracy is s. Were one concerned with

---

<sup>1</sup> When we study the decay of Y to some criterion--decay function asymptote, relative loss, absolute loss, etc.--as a function of amount of training, D becomes the dependent variable, a measure which is analogous to Hull's  $n$ . In one retention paradigm,  $\underline{D} = f(\underline{X})$ ; in another,  $\underline{Y} = f(\underline{D})$ .

removing effects of chance probability of correct responses, accuracy would become a function of  $\underline{s}$  and of the chance probability of correct responses inherent in the test task. We use the terms  $\underline{s}$  and accuracy interchangeably here.

The dimension of  $\underline{t}$  reflects rate of performance. In the realms of commerce, few indeed are the tasks wherein some sort of time pressure does not occur. Whereas only one way exists to define accuracy if we ignore the question of possible correction for chance-correct responses, rate can be defined in various ways. Among these are  $\underline{s}/\underline{t}$ , where  $\underline{s}$  is score;  $\underline{r}/\underline{t}$ , where  $\underline{r}$  is proportion of test items responded to, whether or not correctly; and  $\underline{n}/\underline{t}$ , where  $\underline{n}$  equals unity.<sup>1, 2</sup>

$\underline{Y}_{\underline{D}} = 0$  has two dimensions. Specification in kind depends upon selection of a rate definition. We choose  $\underline{s}/\underline{t}$ , although with some misgivings. Hence, the dimensions of  $\underline{Y}_{\underline{D}} = 0$  will be taken as  $\underline{s}$  and  $\underline{s}/\underline{t}$ .

The dimensions of  $\underline{Y}$  when  $\underline{D}$  is a dependent variable will be  $\underline{s}$ ,  $\underline{s}/\underline{t}$ , and  $\underline{D}$ , since a definition of  $\underline{D}$  as number of days to extinction or to the asymptote of a decay function does not appear to be very meaningful out of the context of the value  $\underline{Y}_{\underline{D}} = 0$ .

---

<sup>1</sup> The consumer of skill might be indifferent to all rates in excess of some minimum. Under such conditions, rate would be increasingly favorably received until the minimum acceptable rate was attained. Thereafter, improvement in rate would not reflect better performance in a domain of practicality.

<sup>2</sup> Rate may change with warmup, fatigue, and other factors associated with extent of test. Let us consider our rate measure as a mean value of rate for negotiation of a test whose length and difficulty are sufficiently great to minimize warmup and sufficiently curtailed to minimize fatigue.

It will not be missed by the reader that s is analogous to Hull's p, that t is analogous to his  $s_t$ , or that D is analogous to his n.

This completes the specification of dimensions of Y in number and kind. There remains to be specified the relationships between them. One objective of the section which follows will be specification of these relationships.

III. RELATIONS BETWEEN  $\bar{X}$ ,  $s$ ,  $s/t$ , AND  $D$ :  
A PRELIMINARY MODEL FOR STUDIES IN EDUCATION

A. Why Model Now?

Let us confine ourselves for the moment to the situation in which  $D = 0$ , or at least appreciably near to 0. Two general designs may be distinguished:

Design A:  $\bar{X}_1, t_1$

Design B:  $t_0, \bar{X}_1, t_1, \bar{X}_2, t_2, \dots, \bar{X}_n, t_n$

We might call Design A the poor man's design. In it, we train once and test once. It is a fine design for pilot work, including work directed at scanning potential predictor variables to see if one cannot be found which will permit operating on the basis of quite small treatment  $N$ 's. For each student in Design A research, we obtain a single point falling on a function which he would yield if trained and tested under Design B conditions. The trouble is that different students in the same treatment group will present different points with respect to  $\bar{X}$ . Design A measures require extrapolation to a common value of  $\bar{X}$ . Linear extrapolation is easily done. If one rejects linear extrapolation as I do, then an alternative model is required before Design A measures can be summarized.

What it comes down to is that Design A requires specification of relations of  $\bar{X}$ ,  $s$ , and  $s/t$  as a prerequisite to extrapolation of the performances of diverse individuals and of diverse groups to some common point on the abscissa, whereas Design B permits these relations to be settled upon after the fact of data collection.

Hence, in the model to be developed, I plan to take myself seriously from a standpoint of the requirements of programmed Design A investigations. From a standpoint of Design B investigations, I expect the model to show itself in need of revision with its first test.

In paragraph B, empirical measures will be defined and their theoretical relations preliminarily asserted. In paragraph C, all terms in the theoretical expressions will be defined in relation to their empirical requirement and evaluative procedure. In paragraph D, steps underlying generation of the function Y of X, where D = 0, will be set down. In paragraph E, a stab will be taken at the function Y of X, where D = 0 to n.

**B. Terms and Relations**

Empirical terms and preliminary assumptions concerning their theoretical relations are given below. The bar below a term indicates it to be an empirical measure or summary Y value based on empirical measures.

Empirical (Obtained) Measures

1.  $\bar{X}_n$  = training time, nth training trial

Theoretical (Assumed) Relations

1.  $X_n = \frac{\bar{X}_1 - k}{n^a} + k$

where:  $\bar{X}_1$  = curve-fitted value of training time, 1st training trial<sup>1</sup>

---

<sup>1</sup> The symbol  $\bar{X}$  is used to designate fitted values of the function  $X$  of  $n$ .

Empirical (Obtained) Measures

Theoretical (Assumed) Relations

k = limiting value of  
the function X of n

n = nth training trial,  
where a trial is one  
negotiation of a  
Content A

a = slope constant

2.  $s_n$  = proportion of test items--  
or other test task--cor-  
rectly responded to, when  
student is tested following  
training on the nth trial

2.  $s_n = 1 - e^{-b(C + T)}$   
where: C = hypothetical training  
time,<sup>1</sup> whose cuset is  
Y = 0 and terminus is  
T = 0, and where T is  
real training time

$$T = \sum_{i=1}^n X$$

b = slope constant

e = base, natural logarithms

3.  $t_n$  = testing time, when student  
is tested following nth  
training trial

3.  $t_n = ms_n X_n$   
where: m = slope constant

---

<sup>1</sup> The construct C permits dealing with the case where Y is greater than 0 at T = 0.

Although we do not show it as part of the set of empirical measures of the model, a later view on rate--one which is more representative--might require us to include the measure  $\underline{r}_n$  (proportion of test items overtly--whether correctly or incorrectly--responded to, when student is tested following training on the nth trial). Its theoretical analogue would be an exponential growth function identical in form to that for  $\underline{s}_n$ ; that is:  $\underline{r}_n = 1 - e^{-h(C + T)}$ , where h is a slope constant.

We have misgivings about the equation for  $\underline{t}_n$ . At a later date, we may shift to such an alternative as  $\underline{t}_n = u r_n X_n$ , where u is a slope constant, or even to  $\underline{t}_n = m X_n$ , or to some other alternative which the data may indicate.

Earlier we stated, with misgivings, that we would presently accept  $\underline{s}/\underline{t}$  as our rate measure. Hence, it remains for us to relate  $\underline{s}$  and  $\underline{s}/\underline{t}$ . Below we do this.

<u>Relation between Empirical Measures</u>	<u>Theoretical Analogue to Relation between Empirical Measures</u>
$\underline{Y}_n = \underline{s}_n (\underline{s}_n / \underline{t}_n) = \underline{s}_n^2 / \underline{t}_n$	$\underline{Y}_n = \underline{s}_n^2 / \underline{t}_n = \underline{s}_n^2 / m \underline{s}_n X_n = \underline{s}_n / m X_n$

Thus, we have settled, at this time, upon a multiplicative relationship between accuracy and rate. Had we defined rate as  $\underline{r}/\underline{t}$ , the relationship would have been  $\underline{Y}_n = \underline{s}_n \underline{r}_n / \underline{t}_n$ , with corresponding change in the theoretical analogue. There seems these days to be a good deal of reservation concerning employment of multiplicative relationships. Let us look quickly at some alternatives. By far the simplest ones are arithmetical ones. In addition to multiplication, we have addition

( $\underline{s} + \underline{s}/\underline{t}$ ), subtraction ( $\underline{s} - \underline{s}/\underline{t}$ , or vice versa), and division ( $\frac{\underline{s}}{\underline{s}/\underline{t}}$ , or  $\underline{t}$ ; or vice versa,  $\frac{\underline{s}/\underline{t}}{\underline{s}}$ , or  $1/\underline{t}$ ). All of these are completely unbelievable as statements of  $\underline{Y}$ . Thus, if one must quarrel with the multiplicative relationship, there is nowhere to go except toward more elegant expressions, which are harder to justify out of the armchair. Examples of such alternatives are:

Empirical

$$\underline{Y}_n = \left( \frac{\underline{s}_n}{\underline{t}_n} \right)^{-\underline{s}_n}$$

Theoretical

$$\underline{Y}_n = \left( \frac{\underline{s}_n}{\underline{t}_n} \right)^{-\underline{s}_n} = \left( \frac{1}{m\underline{X}_n} \right)^{-\underline{s}_n}$$

The possibilities are very many indeed. I feel that the multiplicative relationship is a believable one with which we can live until the weight of evidence points the direction toward change.

Modification of the model will be indicated when it becomes evident that the theoretical function  $\underline{Y}$  of  $\underline{X}$  does not follow closely the function  $\underline{Y}$  of  $\underline{X}$ . The model can only be tested as a consequence of Design B investigations.

It should be noted that Design A extrapolations must be premised on the assumption that  $C = 0$ . The reason for this, discussed below, is that the value of  $C$  cannot be established on the basis of data afforded by Design A. Assuming that in practice  $C$  will always equal some value in excess of 0, Design A extrapolations will consistently overestimate the true value of  $\underline{Y}$  at the value of  $\underline{X}$  to which the extrapolation is made if the obtained value of  $\underline{X}$  is under the value to

which extrapolation occurs. Conversely, they will consistently underestimate the true value of Y if the obtained value of X is over the value to which extrapolation occurs.

C. Evaluation of Empirical Constants k, C, a, b, m

One purpose of this paper is to determine, given a model such as is proposed here, the empirical requirement underlying assessment of the model. All critical unknowns of the proposed model are constants whose value may be appraised on the basis of certain empirical measures. Below, for each of these constants, the empirical requirement and evaluative procedure is presented. The subscript 0, applied to s and t, refers to test measures on a pre-training test. Other subscripts are self-explanatory. The procedure for estimating k from data is taken from Lewis (5, p. 56).<sup>1</sup>

<u>Constant</u>	<u>Empirical Requirement</u>	<u>Procedure</u>
k	$\underline{X}_1, \underline{X}_2, \underline{X}_3, \underline{X}_4, \underline{X}_5$	1. Plot $\underline{X}$ against n. Fit a smooth curve to the data by inspection. (In this plot, $\underline{X}$ appears on the Y-axis, n on the X-axis.)

<sup>1</sup> An undesirable characteristic of this procedure is to be noted in its application by Lewis (5, p. 61). His freehand fit of values plotted in his Figure 4.5 (5, p. 60), together with extrapolation of Y-values, yields an asymptote of -.29, which--as Lewis notes-- is impossible. My result, using the same data, is an asymptotic value of .13.

Constant      Empirical Requirement

Procedure

2. Choose two points--( $n.$ ,  $X.$ ) and ( $n..$ ,  $X..$ )--which fall on the smooth curve, which encompass the main curvature, yet which do not fall too near origin or asymptote. Then choose a third point whose value  $n...$  is the geometric mean of  $n.$  and  $n..$  and determine  $X...$  by inspection.

3.  $k = \frac{X.X.. - X...^2}{X. + X.. - 2X...}$

c       $s_0, s_1, \bar{X}_1$

1. Substitute values  $s_0, s_1,$  and  $\bar{X}_1$  into equations:

$$s_0 = 1 - e^{-bC}$$

$$s_1 = 1 - e^{-b(C + \bar{X}_1)}$$

2. Solve simultaneously for C.

a       $k; \bar{X}_1, \bar{X}_2,$  as read  
from smooth curve;

$n = 2$

1. Substitute values  $\bar{X}_1, \bar{X}_2, k,$  and  $n$  into the equation:

$$\bar{X}_2 = \frac{\bar{X}_1 - k}{n^a} + k$$

2. Solve for a.

<u>Constant</u>	<u>Empirical Requirement</u>	<u>Procedure</u>
b	C requirement	1. Substitute values C and $s_0$ into the equation: $s_0 = 1 - e^{-bC}$ 2. Solve for b.
m	$t_1, s_1, \bar{X}_1$	1. Substitute values $t_1, s_1,$ and $\bar{X}_1$ into the equation: $t_1 = ms_1\bar{X}_1$ 2. Solve for m.

A quick review of the empirical requirement associated with evaluation of constants will confirm that only k and a require more data collection than is implied in the following paradigm:

Test Zero - Trial One - Test One

Our concept of a training trial is that of a single negotiation of a content consisting of a collection of items brought together on the basis of conformity to criteria of any one of many classification systems of education. So long as we are willing to study Y strictly on the basis of Trial 1 effects, we need not be concerned with k and a. The foregoing paradigm aims, as education often does, at the big initial increment, rather than at asymptotic performance.

Whether or not consistent with worldly requirements, education seems not in general supported to an extent needed to bring a given performance to asymptote. We hesitate to portray educational requirements appreciably

other than as they are portrayed by educational establishments. Education's more immediate needs for behavioral research would seem better served by a basic paradigm which terminated on a Trial  $n$  rather near to Trial 1 than they would by one rather far out. However, the search for more efficient means for imparting the instruction might be unduly hampered if the search were restricted to the condition that training not transcend Trial 1. When we step into a paradigm involving two or more trials, the constants  $k$  and  $a$  come to interest us.

Whether we need the constant  $a$  at this time is debatable. The values of  $\bar{X}$  defined by the smooth curve employed in evaluation of  $k$  can be used in lieu of those given by the hyperbolic equation which is our initial view of the function  $\bar{X}$  of  $n$ .

Assuming that  $k$  represents a physiological limit in speed of reaction, then for a given form of test performance, there will be one and only one limit  $k$ . The student will vary in his rate of approach to  $k$ --as evidenced by shifts in the value of  $a$ --with variation in conditions of acquisition. However,  $k$  will be unaffected. Hence, once  $k$  has been evaluated with respect to a given content and test form, its value can be assumed for all consequent investigations bearing upon that content and test form. At any rate, we will assume so for the present.

However, whatever procedure we employ to determine slope of the function  $\bar{X}$  of  $n$ , it will be necessary to obtain measures for several consecutive training trials. We have used five trials here, since this many trials yield a sufficient number of points to permit curve-fitting by inspection with some little confidence in the fit.

Hence, where  $D = 0$ , we may distinguish three different paradigms which will be employed in various stages of research:

Design A: Train-Test 1.

Design B': Test 0, Train-Test 1.

Design B'': Test 0, Train-Test 1...Train-Test 5.

D. Steps in Generation of  $Y = [1 - e^{-b(C + T)}]^{1/mX}$

Below, in seven steps, the theoretical function  $Y$  of  $X$  is generated from the following data:  $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_5, \underline{s}_0, \underline{s}_1, \underline{t}_1$ . The procedure is illustrated by means of two examples, which share common values of  $\bar{X}_1, \underline{s}_1$ , and  $\underline{t}_1$ , but differ in values of  $\underline{s}_0, \bar{X}_2, \dots, \bar{X}_5$ .

Step 1. Evaluation of  $k$

Discussion: The first step in obtaining  $k$  from empirical data is to fit a curve to the data by inspection. The assumption underlying the procedure employed to obtain  $k$  is that  $\underline{X}$  is a simple hyperbolic function of  $n$ .

Procedure:

Given:  $\underline{X}_1 = 2.05$   
 $\underline{X}_2 = 1.27$   
 $\underline{X}_3 = 1.23$   
 $\underline{X}_4 = 1.14$   
 $\underline{X}_5 = 1.13$

Given:  $\underline{X}_1 = 2.03$   
 $\underline{X}_2 = 1.24$   
 $\underline{X}_3 = 1.15$   
 $\underline{X}_4 = 1.13$   
 $\underline{X}_5 = 1.12$

1. Fit curves by inspection:

$$\bar{X}_1 = 2.00$$

$$\bar{X}_1 = 2.00$$

$$\bar{X}_2 = 1.32$$

$$\bar{X}_2 = 1.26$$

$$\bar{X}_3 = 1.20$$

$$\bar{X}_3 = 1.16$$

$$\bar{X}_4 = 1.16$$

$$\bar{X}_4 = 1.13$$

$$\bar{X}_5 = 1.14$$

$$\bar{X}_5 = 1.12$$

2. Choose two points which fall on the fitted curve:

$$\bar{X}_1 = n., X. = 2.00$$

$$\bar{X}_1 = n., X. = 2.00$$

$$\bar{X}_5 = n.., X.. = 1.14$$

$$\bar{X}_5 = n.., X.. = 1.12$$

3. Choose a third point  $X...$  such that  $n...$  is the geometric mean of  $n.$  and  $n..$ :

$$a. \sqrt{n.n..} = \sqrt{5} = 2.236 = n... \quad \sqrt{n.n..} = \sqrt{5} = 2.236 = n...$$

$$b. X... = 1.28$$

$$X... = 1.22$$

4. Solve for  $k$  in the equation:

$$k = \frac{X.X.. - X...^2}{X. + X.. - 2X...}$$

$$k = 1.10$$

$$k = 1.10$$

### Step 2. Evaluation of a

**Discussion:** The relation between length of training interval on any trial in a series ( $X$ ) and the trial position in the series ( $n$ ) is given by the equation:

$$\bar{X}_n = \frac{\bar{X}_1 - k}{n^a} + k$$

where:  $n$  = nth training trial

$\bar{X}$  = training time in minutes on any trial

$a$  = slope constant

$k$  = asymptote of the function  $X$  of  $n$

Given  $k$ ,  $\bar{X}_1$ , and  $\bar{X}_2$ ,  $a$  is obtained straightforwardly.

Procedure:

Given:  $k = 1.10$

$\bar{X}_1 = 2.00$

$\bar{X}_2 = 1.32$

Given:  $k = 1.10$

$\bar{X}_1 = 2.00$

$\bar{X}_2 = 1.26$

1. Substitute given values into the equation and rearrange terms:

$$n^a = (\bar{X}_1 - k)/(\bar{X}_2 - k) \qquad n^a = (\bar{X}_1 - k)/(\bar{X}_2 - k)$$

$$2^a = (2 - 1.1)/(1.32 - 1.1) \qquad 2^a = (2 - 1.1)/(1.26 - 1.1)$$

$$= .9/.22$$

$$= 4.091$$

$$= .9/.16$$

$$= 5.625$$

2. Solve for  $a$ :

$$a \log 2 = \log 4.091$$

$$a = \log 4.091/\log 2$$

$$= .61183/.30103 = 2.03$$

$$= 2.0 \text{ (rounded)}$$

$$a \log 2 = \log 5.625$$

$$a = \log 5.625/\log 2$$

$$= .75012/.30103 = 2.49$$

$$= 2.5 \text{ (rounded)}$$

### Step 3. Evaluation of C and b

**Discussion:** The concept C is necessary where an absolute zero  $\underline{Y}$ -value does not characterize the student at the onset of formal training and we wish to view  $\underline{Y}$  as an exponential function of training time, including both formal and pre-formal components. C can possibly be dispensed with if  $\underline{Y}$  is viewed as some other kind of function of  $\underline{X}$ .

#### **Procedure:**

Given:  $s_0 = .6$

$s_1 = .8$

$\bar{X}_1 = 2$

Given:  $s_0 = .2$

$s_1 = .8$

$\bar{X}_1 = 2$

1. Determine value of the exponent of  $s_n = 1 - e^{-b(C + T)}$ .

$$.6 = 1 - e^{-bC} = 1 - e^{-.916}$$

$$.2 = 1 - e^{-bC} = 1 - e^{-.223}$$

$$.8 = 1 - e^{-b(C+2)} = 1 - e^{-1.609}$$

$$.8 = 1 - e^{-b(C+2)} = 1 - e^{-1.609}$$

2. Solve simultaneously for C:

$$C = 2.64$$

$$C = .322$$

3. Substitute the value of C into exponents and determine b, the slope constant:

$$b = .347$$

$$b = .693$$

### Step 4. Generation of the Function X of n

**Discussion:** Values of X enter into theoretical determination of Y. These values are obtained by successive application of the equation of Step 2, where  $\bar{X}_1$ , k, and a are given.

Procedure:

Given:  $\bar{X}_1 = 2$

$k = 1.1$

$a = 2$

Given:  $\bar{X}_1 = 2$

$k = 1.1$

$a = 2.5$

1. Solve for  $X_2$ :

$$X_2 = (2 - 1.1)/2^2 + 1.1$$

$$= .9/4 + 1.1$$

$$= .225 + 1.1 = 1.325$$

$$= 1.32 \text{ (rounded)}$$

$$X_2 = (2 - 1.1)/2^{2.5} + 1.1$$

$$= .9/5.657 + 1.1$$

$$= .159 + 1.1 = 1.259$$

$$= 1.26 \text{ (rounded)}$$

2. Solve for  $X_3, X_4, X_5$ . Values are:

$$X_3 = 1.20$$

$$X_4 = 1.16$$

$$X_5 = 1.14$$

$$X_3 = 1.16$$

$$X_4 = 1.13$$

$$X_5 = 1.12$$

Lo and behold, theoretical values thus obtained are identical to inspection-fitted curve values given in Step 1. This sort of coincidence between empirical and theoretical is best sought out in illustrative material such as this. The empirical domain is less obliging.

Step 5. Generation of the Function  $s$  of  $(C + T)$

Discussion: Values of  $s$  enter into theoretical determination of  $Y$ .

These values are obtained by successive application of the equation:

$$s_n = 1 - e^{-b(C + T)}$$

where:  $s$  = proportion of test items correctly answered

$n$  = nth training trial

$e$  = base, natural logarithms

$b$  = slope constant

$C$  = pre-formal (hypothetical) training time

$T = \Sigma X$

Below, we solve through  $s_3$ .

Procedure:

Given:  $b = .347$

$C = 2.64$

$T_2 = 3.32 = (X_1 + X_2)$

$T_3 = 4.52 = (X_1 + X_2 + X_3)$

Given:  $b = .693$

$C = .322$

$T_2 = 3.26 = (X_1 + X_2)$

$T_3 = 4.42 = (X_1 + X_2 + X_3)$

1. Substitute:

$s_2 = 1 - e^{-.347(2.64 + 3.32)}$

$s_3 = 1 - e^{-.347(2.64 + 4.52)}$

$s_2 = 1 - e^{-.693(.322 + 3.26)}$

$s_3 = 1 - e^{-.693(.322 + 4.42)}$

2. Solve for  $s$ :

$s_2 = .874$

$s_3 = .916$

$s_2 = .916$

$s_3 = .962$

Step 6. Evaluation of  $m$

Discussion: Only the slope constant  $m$  of the equation  $t_n = m s_n X_n$  remains to be determined before theoretical values of

the function Y of X can be generated. Since  $t_1$ ,  $s_1$ , and  $\bar{X}_1$  are the same for both examples, the value of m will be the same.

**Procedure:**

Given:  $t_1 = .64$

$s_1 = .8$

$\bar{X}_1 = 2$

1. Substitute given values into the equation  $t_1 = ms_1\bar{X}_1$ :

$$.64 = m(.8)(2)$$

2. Solve for m:

$$m = .4$$

**Step 7. Generation of the Function  $Y = [1 - e^{-b(C + T)}] / mX$**

**Discussion:** We now possess all the information required to generate theoretical values of the function Y of X. These expected functions will be accepted to the extent that they coincide with functions based on the equation  $\underline{Y} = \underline{s}^2/\underline{t}$ . The functions are plotted in Figure 1.

**Procedure:**

Given:  $s_0 = .6$

$s_1 = .8$

$s_2 = .874$

Given:  $s_0 = .2$

$s_1 = .8$

$s_2 = .916$

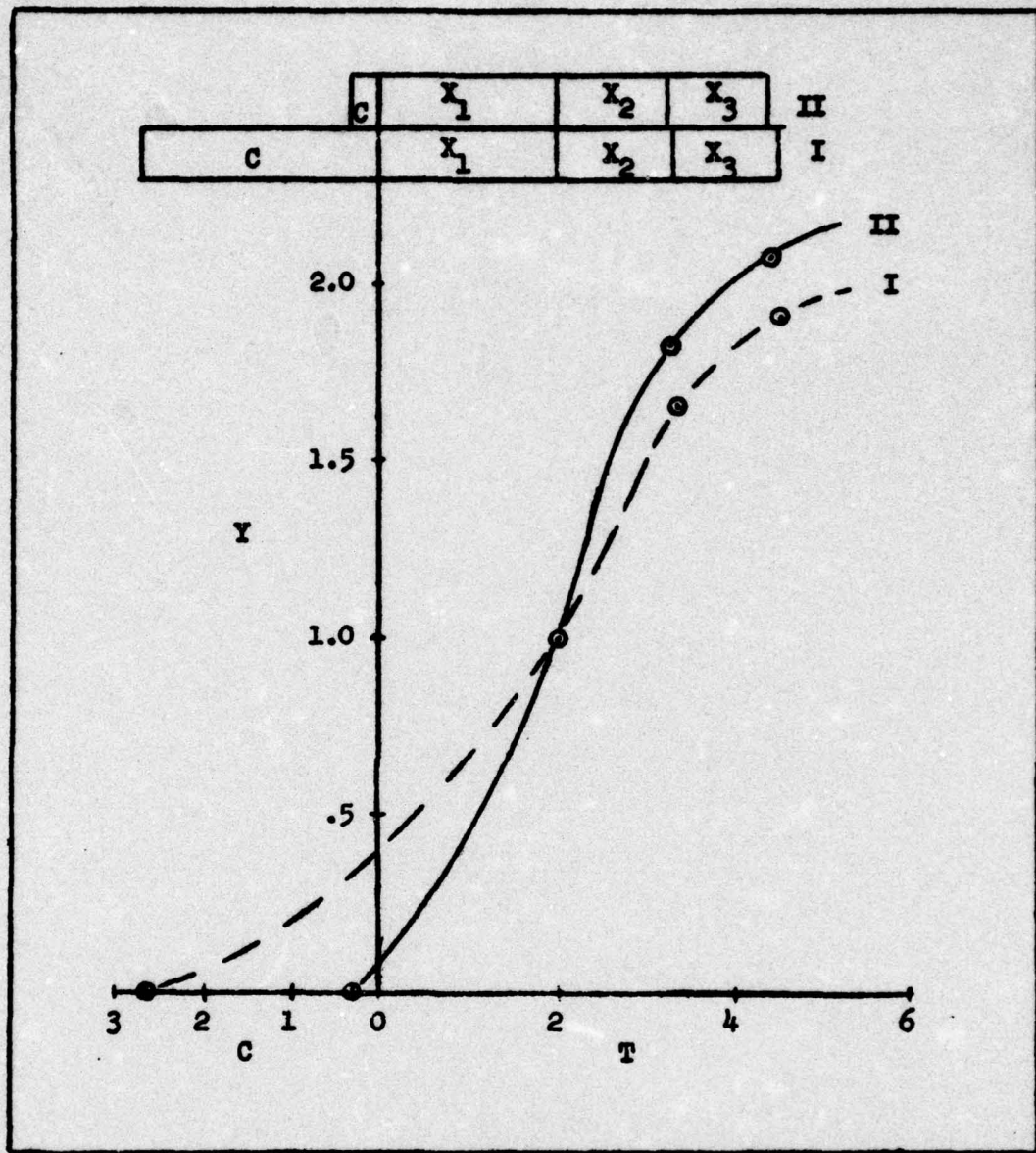


Figure 1.  $Y = [1 - e^{-b(C + T)}] / mX.$

$$s_3 = .916$$

$$c = 2.64$$

$$n = .4$$

$$\bar{x}_1 = 2.00 \text{ (fitted value)}$$

$$x_2 = 1.32$$

$$x_3 = 1.20$$

$$s_3 = .962$$

$$c = .322$$

$$n = .4$$

$$\bar{x}_1 = 2.00 \text{ (fitted value)}$$

$$x_2 = 1.26$$

$$x_3 = 1.16$$

1. Substitute:

$$y_1 = .8/.4(2)$$

$$y_2 = .874/.4(1.32)$$

$$y_3 = .916/.4(1.20)$$

$$y_1 = .8/.4(2)$$

$$y_2 = .916/.4(1.26)$$

$$y_3 = .962/.4(1.16)$$

2. Solve for Y:

$$Y = 0 \text{ at } T = -2.64$$

$$Y = 1.00 \text{ at } T = 2.00$$

$$Y = 1.66 \text{ at } T = 3.32$$

$$Y = 1.91 \text{ at } T = 4.52$$

$$Y = 0 \text{ at } T = -.322$$

$$Y = 1.00 \text{ at } T = 2.00$$

$$Y = 1.82 \text{ at } T = 3.26$$

$$Y = 2.07 \text{ at } T = 4.42$$

B. Retention Over Time (D ≠ 0)

Casual examination of literature supports the general notion that meaningful material which is rather well learned and then retention-tested by the single recall method will yield performance which asymptotes about a month after conclusion of training, at a rather high level of retention. Curves will be negatively accelerated, descending.

Our concern with retention is one of response definition. Earlier, I noted that a representative design for education cannot accept a definition of response which fixes its position as temporally contiguous with training events. It is conceivable that performance associated with Training Treatment A will be much better for the same investment than that associated with Training Treatment B if the two are assessed just at the end of training, but that Treatment B performance will emerge as the superior one when assessment occurs over time. Assuming both treatment groups go to the same residual retention value, we are interested in the number of days it takes in the case of either group; or, assuming they go to different residual asymptotes, we are interested in the relative heights of these asymptotes.

Since we have gotten into the habit of talking in terms of theoretical and obtained functions, we will play the game one more time for  $D \neq 0$ . But first let us correct an impression of certain investigators of retention--that the interval  $D$  can be made contentless by an act of definition. It cannot. A content  $I$  will always be present. The job of dimensionalizing the universe of  $I$  will not be attempted here. Rather, we will "control"  $I$  at the level of bedlam which characterizes it in the life of any student.

Wherever training is concluded, the value  $\underline{Y}$  is taken for one sub-group of a treatment group. Presumably, the sub-group is quite representative of the group as a whole and  $\underline{Y}$  closely approximates  $Y$ .

For the present, we do not much care where  $Y$  is on the acquisition function, so long as it is somewhat above its value at  $T = 0$ . We assume the following theoretical relationships:

1.  $s' = e^{-b'D}(s - v) + v$

where:

$s'$  = test score when tested the first time following completion of training, where the interval between training and testing is  $D$  days

$s$  = test score at  $D = 0$ , taken on a group or individual considered in every relevant way equal to that from which  $s'$  is obtained

$v$  = the value to which the function  $s'$  of  $D$  is asymptotic; that is, the value which is closely approached in  $n$  days

$b'$  = slope constant

$e$  = base, natural logarithms

2.  $t' = e^{-m'D}(w - t) + t$

where:

$t'$  = testing time when tested the first time following completion

of training, where the interval between training and testing is D days

t = testing time at D = 0

w = the value to which the function t' of D is asymptotic; that is, the value which is closely approached in n days

m' = slope constant

e = base, natural logarithms

$$3. Y' = s'^2/t'$$

where:

$$Y' = Y \text{ at } D = 0$$

$$Y' = v^2/w + t \text{ at } D = \text{infinity}$$

The obtained function  $\underline{Y}'$  of  $\underline{D}$  is, of course,  $\underline{Y}' = \underline{s}'^2/\underline{t}'$ .

We are not yet ready to treat retention designs with the thoroughness they deserve. Their more elaborate treatment will be postponed to a date when a model for initial acquisition is at hand which has been tested and found acceptable for the times.

#### IV. SUMMARY

Where behavioral research aspires to serve education's more immediate teaching needs, laboratory performances must be representative of those which education is charged to teach. Else, laboratory findings may not apply to the educational domain. A laboratory performance which is representative of utilitarian performances of the world of commerce will usually be multidimensional in the sense that its adequate description presupposes initial assessment along two or more dimensions. An adequate description of multidimensional performances presupposes (1) enumeration of those underlying dimensions which meet worldly criteria of necessity and joint worldly-laboratory criteria of sufficiency, and (2) statement of functional relations holding between these dimensions, so that multidimensional performance can be unidimensionally summarized.

Beyond the dictum of representativeness, the notion has been advanced here that it is an improper view of aptitude mechanics to hold that one's particular aptitudes are consequential to acquisition but inconsequential to demonstration of acquisition. We have challenged the procedure whereby the individual is allowed to pace himself during acquisition but not during testing.

An acquisition model for education--consistent with representativeness and the self-paced test--was presented. Preliminary views on a model for retention were also set forth, although in less detail and more tentatively than the model for acquisition.

## REFERENCES

1. Brunswik, Egon. Systematic and representative design of psychological experiments, with results in physical and social perception. Berkeley: University of California Press, 1947.
2. Estes, William K., et al. Modern learning theory. New York: Appleton-Century-Crofts, 1954.
3. Guilford, J. P. Psychometric methods. (2d ed.) New York: McGraw-Hill, 1954.
4. Hull, Clark L. Principles of behavior: an introduction to behavior theory. New York: Appleton-Century-Crofts, 1943.
5. Lewis, Don. Quantitative methods in psychology. New York: McGraw-Hill, 1960.
6. Skinner, B. F. The behavior of organisms: an experimental analysis. New York: Appleton-Century-Crofts, 1938.

## AFTERTHOUGHT

Examination of Y will reveal a more complex concept of performance than characterizes more explicit theory-builders of the present. This does not mean to me that we should turn our backs on their efforts. On the contrary, we will surely gain in understanding of problems attending conceptualization of units of performance by examining explicit treatments of others. One of the best illustrations is provided by Estes,<sup>1</sup> who has developed one of the more rigorous models of behavior.

Estes' acquisition situations differ critically from our own. His situations feature a response A--occurring under free responding conditions--of zero duration. The passage of time in Estes' system occurs entirely under conditions of occurrence of response B (all behaviors not A). The measure of A comes down to enumerating A and timing B.

Perhaps the best place to begin in describing that part of Estes' system which concerns us here is with p. The intervening variable p (response probability) is the output plug of the system's meager theoretical superstructure. To it are tied those measures which serve as dependent variables at an empirical level of discourse. The dependent variable tied to p may have either rate or accuracy characteristics, but not both, since the systemic situation precludes a response A having duration.

---

<sup>1</sup> Estes, William K. The statistical approach to learning theory. In Koch, Sigmund (Ed.). Psychology: a study of a science. Volume II, General systematic formulations, learning, and special processes. New York: McGraw-Hill, 1959.

Examples of "rate type" dependent variables in Estes' system are L (latency of A, or duration of B), which equals  $h/p$ , and  $r$  (average rate at which response A is made;  $r_n$ , after  $n-1$  reinforcements), which equals  $p/h$ , where  $h$  is the asymptotic value of exposure to the independent variable prior to occurrence of response A (asymptotic trial latency).

An "accuracy type" dependent variable is  $e(K)$ , the number of errors made in  $K$  trials. Before elaborating on  $e(K)$ , we will further define  $p$  in the system.

The output theoretical entity  $p$  is a function of the explicit theoretical input entity  $n$  (number of reinforced trials) and of certain Z-axis variables not yet defined but represented by the parameter (slope constant)  $\theta$ . Unlike many systems, whose slope constant can take on any value, the limits of  $\theta$  are fixed. It may vary from just over 0 to just under or equal to 1. This characteristic contrasts with slope constants in exponential growth functions--for instance,  $b$  of my function  $s$  of training time--which may take on any value.

The following basic equation is currently employed by Estes to relate  $p$  to input:

$$p_n = 1 - (1 - \theta)^{n-1}$$

The "accuracy type" dependent variable  $e(K)$  is defined as the sum of "the values of  $1 - p_n$  for  $n = 1, 2, \dots, K$ ." The sum of these values is given by:

$$e(K) = \frac{1 - (1 - \theta)^K}{\theta}$$

The total number of errors during acquisition equals  $1/\theta$ , since  $(1 - \theta)$  drops almost to zero when  $K$  is large.

Unlike the T-maze, wherein each trial is characterized by a correct or an incorrect turn (equivalent to  $p = 0$ ,  $p = 1$ ), we begin with a 20-item test, such that we can obtain an accuracy equivalent of  $p$  on a single test. We need not resort to  $e(K)$ . By treating  $s$  as a proportion, rather than a number  $c$  of correct responses to a test of  $I$  length, we have a scale of  $s$  which has the same limits as the scale of  $p$ . Bearing in mind that  $s$  is conditional on a test length  $I$ ,  $s_n$  can be treated equivalently to  $p_n$ . However, just as we used the concept  $C$  (hypothetical training time) to account for the fact that the student does not come to us at  $Y = 0$  ( $s = 0$ ,  $p = 0$ ), we must scratch for a value  $Q$ --a hypothetical number of trials antedating observation of the student--which augments  $n$ .

In the body of the paper, we assumed  $X$  (trial training time) to be a hyperbolic function of  $n$  (number of trials). If so, then the limiting value of  $n$  at which  $X$  is infinite is a true origin of the scale of  $n$ . The value  $n = 1$ , based on a first trial value, is not the true origin of the function  $X$  of  $n$ ; neither is it one unit to the right of a true origin. Our units of  $n$  are arbitrary ones which must be rectified through use of the value  $Q$ .

We begin with a scale of  $n$  such that  $n = 1$  reflects Trial 1 in the experimental situation;  $n = 2$ , Trial 2; etc. Placing  $n$  on the ordinate and  $X$  on the abscissa, we determine the asymptotic value  $k'$  of  $n$ . Let us say a value of  $.7$  is obtained. Hence,  $X$  is infinite at  $.7 n$ , a true origin.

Just as the rat in the Skinner box does not forsake forever the proffered bar, our student does not take infinite time on Trial 1. In fact, he is usually well along on his function  $X$  of  $n$  by the time of Trial 1, as indicated by extrapolation to  $k'$ . Where  $k' = .7$ , we may say that .3 trials have occurred by the time Trial 1 occurs. That is,  $Q = X_1 - k' = .3$ . So the proper trial value after one experimental trial is 1.3. Without getting into either what is meant by .3 of a reinforcement or what constitutes a reinforcing event in our situation, we assert that  $n - 1 = .3$  at Trial 1. Let us say the slope constant  $\theta$  equals .5. Hence, if we use Estes' relation (modified to include  $Q$ ) for the development of  $s$ , we have:

At origin:

$$s_0 = 1 - (1 - \theta)^0 = 1 - 1 = 0.$$

At completion of Trial 0 ( $Q$  training interlude):

$$s_{0+Q} = 1 - (1 - \theta)^Q = 1 - (1 - .5)^3 = .188$$

At completion of Trial 1:

$$s_{1+Q} = 1 - (1 - \theta)^{Q+n} = 1 - (1 - .5)^{1.3} = .594$$

We use  $n$  in lieu of  $n - 1$  on the assumption that the reinforcing event of a given trial occurs during the acquisition phase of the trial and so will have its effect upon test performance of that trial. It is worth a paper several times the length of this one to delve into a description of reinforcers and the conditions of their presentation during the learning sessions of education. One reason for selection

of elapsed training time in lieu of reinforced trials in conjunction with writing an initial expression of the relation of  $s$  to the X-axis was the great uncertainty which surrounds the concept of reinforcers in its application to education.

I began by asserting that Estes' concept of performance differed from one which I claim represents the educational requirement. This is evident when we compare my rate measure  $s/t$  with his rate measure  $p/h$ . My  $k$  is analogous to his  $h$ , and my  $s$  similar to his  $p$ . Yet I cannot accept  $s/k$  as a rate measure. If, as is generally supposed, rate of acquisition is independent of asymptote, then what is important to me in the expression  $s/t$  is the rate at which the student approaches ( $s = 1.0$ )/( $t = k$ ). An asymptotic expression in the denominator of the rate ratio transforms the concept of rate into a linear function of  $s$  ( $rm = s$ ).

If the conceptualizers of educational research cannot, at the moment, fit right into someone else's model of performance or of stimulus-response relations, still there is something to be gained by examination of the way in which the more rigorous systematists go about modelling performance and its more fundamental relations to input.