ALL-OR-NONE SUBPROCESSES
IN THE LEARNING OF COMPLEX SEQUENCES

by

Robert A. Bjork

TECHNICAL REPORT NO. 98

June 22, 1966

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

ALL-OR-NONE SUBPROCESSES

IN THE LEARNING OF COMPLEX SEQUENCES

Robert A. Bjork*

Stanford University

Abstract

This paper reports a study designed to investigate whether the all-or-none conception of the learning process can be extended to a learning task more complex than conditioning or simple verbal association. The experimental task is to learn numerical sequences by anticipating each new member of the sequence. Although the obtained sequence learning appears very complex, it proves to be analyzable into constituent all-or-none subprocesses.

---

*Now at the University of Michigan. This work was accomplished at Stanford University while the author held, during the initial stages, a National Defense Education Act Graduate Fellowship and, during the later stages, a National Science Foundation Terminal Graduate Fellowship. Support for the research was provided, in part, by grants NSF 6-24264 and MH 6154 from the National Science Foundation and Institutes of Mental Health and, in part, by Contract Nonr 225(73) between the Office of Naval Research and Stanford University. The author is indebted to Professor W. K. Estes for his encouragement and counsel during all stages of the study.

The success of specific models for conditioning and verbal association that embody the all-or-none conception of the learning process has led to speculation (Estes, 1964; Restle, 1964) that it might be profitable to extend the all-or-none conception to more complex learning. Although the simplest all-or-none model, the one-element model, which assumes a single all-or-none learning step is obviously inadequate for situations more complex than conditioning or simple association, it may be the case that complex learning consists of a number of successive and/or simultaneous all-or-none processes. There is some evidence for such a position. Bower and Theios (1964) describe avoidance behavior well with a model based on the assumption that learning the avoidance response involves two successive all-or-none subprocesses. Also, Restle (1964) interprets difficult paired associate learning in terms of a multi-stage model in which the stages are all-or-none processes corresponding to association, stimulus discrimination, and response discrimination.

This paper reports a study designed to investigate whether the all-or-none conception can be extended to sequence learning, a learning task considerably more complicated than simple association. The general format of sequence learning is as follows: the subject is presented the first member of a sequence, attempts to predict the second member, is presented the second member of the sequence, attempts to predict the third member, is presented the third member of the sequence, and so on until he achieves an arbitrary learning criterion of N successive correct anticipations. There are two main reasons why the learning of numerical sequences is an appropriate task for this study:

1. As a learning task it is considerably more complex than those tasks which typically exhibit all-or-none learning. In simple associative tasks the stimulus and associated "correct" response are constant. In sequence learning the stimulus presented continually changes and the response designated as correct also changes. The subject must abstract from the progressing sequence constant relationships upon which to base his prediction of the next sequence member. For example, if the sequence is 0, 2, 1, 3, 2, ..., he must learn that whenever the sequence is incremented by 2 it will then be decremented by 1 . In addition, most sequences involve more than one abstract association; in order to completely learn the sequence the subject must abstract several relationships concurrently. Finally, more alternatives are available to the subject in sequence learning than in simple association, and at times these alternatives are equally "reasonable"; much more information must be retained and utilized by the subject.

2. Although the task is considerably more complex, the response protocols generated are formally similar to those resulting from simple association experiments. The subject's responses for a particular sequence can be represented as a trial-by-trial string of errors and successes yielding protocols essentially equivalent to paired-associate or conditioning protocols. This protocol equivalence means that differences and similarities between sequence learning and simple association are maximally

apparent, since the sequence data can be analyzed in a fashion parallel to the standard analyses performed on simple association data.

METHOD

Fifty-nine introductory psychology students at Stanford University participated in this experiment in fulfillment of a course requirement. They were each required to learn twelve numerical sequences which varied considerably in their complexity. All Ss were given the twelve sequences in the same order.

Apparatus A flat black 3 ft. square of plywood was supported vertically on a table, separating the S and the experimenter. In front of the S, mounted on plywood, was a white circle of 12 in. diameter. The ten digits, 0, 1, ..., 9, were printed in black at equal intervals inside the circumference of the circle as shown in Figure 1. Mounted just outside the circle next to each digit was a small light. Any particular digit could be designated by turning on its adjacent light.

The experimenter could control the digit presented on a particular trial by means of a ten-contact rotary switch wired to correspond to the circle of Figure 1. As the switch was turned from one number to another, all numbers in between flashed momentarily as their contacts were touched. The direction of this series of brief flashes enabled the S to tell whether addition or subtraction was being performed. If, for example, one member of a sequence was 2 and the next 7, adding five would light briefly 3, 4, 5, 6, between 2 and 7, whereas subtracting five would light briefly 1, 0, 9, 8, between 2 and 7.

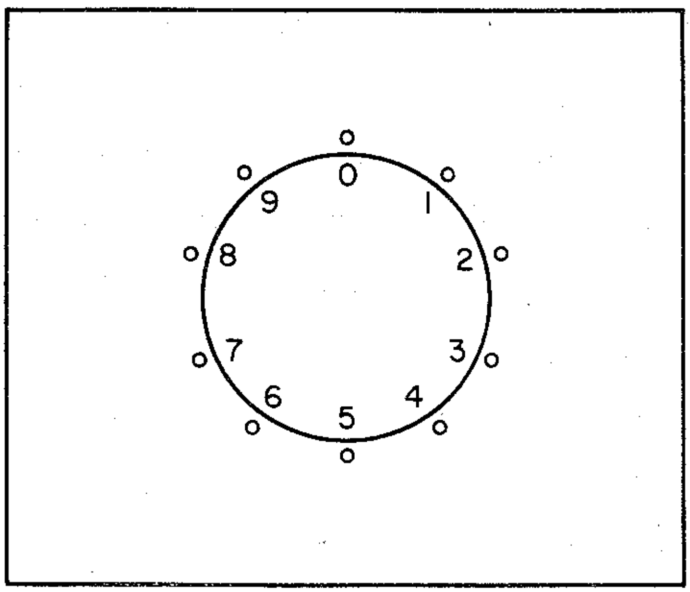

Fig. 1. Display Apparatus. The small circle adjacent to each digit represents a light used to designate that digit.

Instructions *S*s were told that their task would be to discover the scheme or system underlying each numerical sequence. They were also told that the sequences were arithmetic in that each new member of a sequence would be obtained by adding or subtracting some integer to or from the last member of the sequence; it was emphasized that multiplication or division were never involved. Two illustrative sequences were presented and explained to familiarize the *S* with the procedure.

Procedure The method of presentation is diagrammed in Figure 2 where N_0, N_1, N_2, ... represents an arbitrary numerical sequence. In stimulus-response terms, the reinforcement for one trial served as the stimulus for the next trial. Note that a *S* had only the current member of the sequence present and had to remember past members of the sequence.

*S*s responded vocally; the experimenter recorded each response and presented the next member of the sequence. When a *S* anticipated five consecutive sequence members correctly, he was told that he had solved the sequence; there was a short interval; and then the experimenter presented the first member of the next sequence. If a *S* had not started a criterion run of five correct by the twenty-fifth trial, he was told that that was the end of the sequence, and after a short interval, the next sequence was started.

Nature of the sequences The experimental sequences, shown in Figure 3, are generated by operations of two basic kinds. The simpler of the two operations is adding a constant integer. For example, sequence 4: 1, 4, 7, 10, ..., is constructed by adding 3 on every trial. By combining

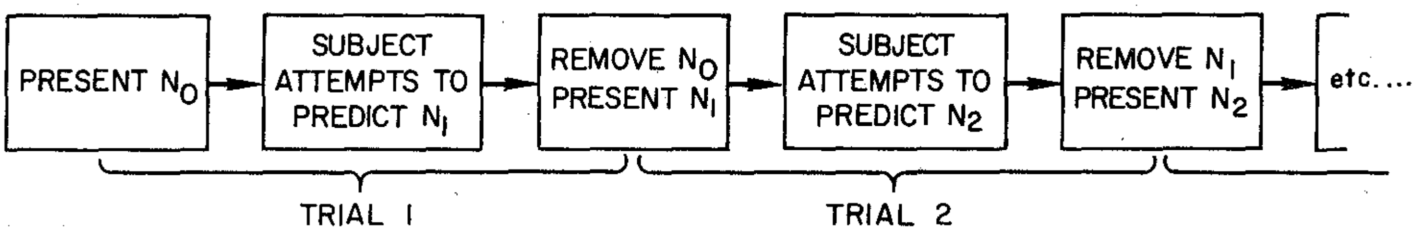

Fig. 2. Method of Presentation. N_0, N_1, N_2, ... represents an arbitrary numerical sequence.

several such operations more complex sequences can be generated; sequence 3: 1, 5, 7, 11, 13, ..., is constructed by alternately adding 4 and adding 2. In this case the period of the operation, add 4, and of the operation, add 2, is 2; i.e., on every second trial one of the operations is applied and on alternate trials the other is applied. In general, then, for $n = 1, 2, 3$, a permissible sequence of this sort can be constructed using n operations each with period n.

The second, and more complicated of the two operations, is an advancing, rather than constant, increment or decrement to the sequence. For example, sequence 1: 1, 2, 4, 7, 11, ..., is formed by adding 1, adding 2, adding 3, etc. Again, more complex sequences can be generated by operations of this kind by using more than one such operation in a single sequence as in sequence 2: 0, 2, 1, 4, 2, ..., constructed by adding 2, subtracting 1, adding 3, subtracting 2, adding 4, subtracting 3, and so on.

Finally, the two operations can be mixed to construct sequences. Sequence 7: 2, 7, 6, 11, 9, 14, ..., is generated by adding 5 on the odd trials and subtracting 1, 2, 3, 4, ... on the even trials

RESULTS

The learning curves obtained for the twelve sequences are shown in Figure 3. These learning curves, to say nothing about more sensitive characteristics of the data, appear too complex to be reflecting any simple model; they do not look like curves arising from any of the typical conceptions of the learning process.

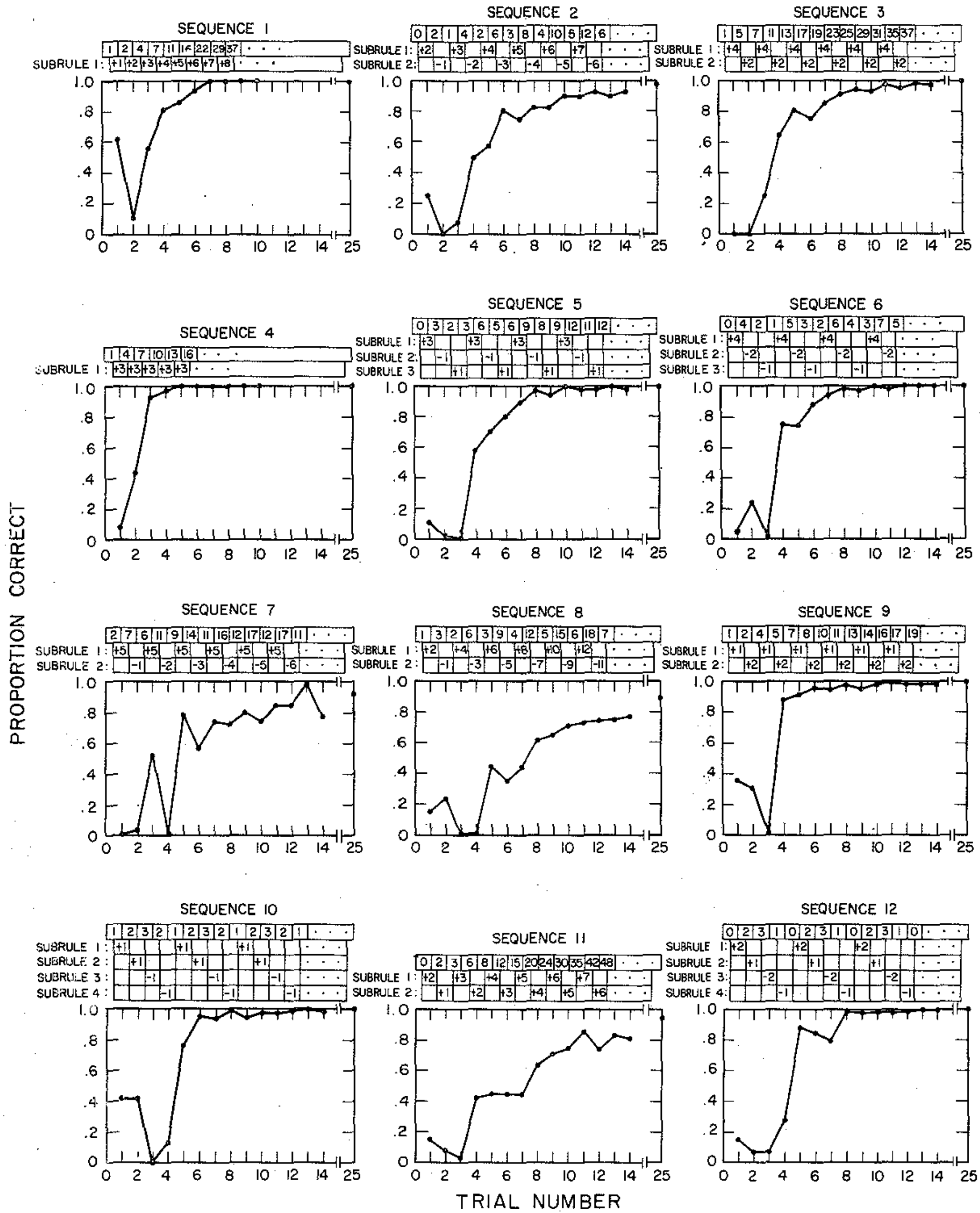

Fig. 3. Learning Curves for Sequences 1-12. Above each learning curve are the actual sequence members together with the subrules used to generate the sequence.

In spite of the general complexity of the sequence data, the possibility is not excluded that subprocesses, i.e., components of the learning defined in some manner, might exhibit all-or-none learning. A process generated by several simple, but interlaced, stochastic subprocesses could appear complex. Even if the component subprocesses are the same, their probabilistic nature and joint generation of the learning process could yield very complex performance.

A subprocess analysis requires a partition of the data made according to some rational breakdown of the learning into components. Consider, in Figure 3, the actual sequences presented. In order to learn any of these sequences a S must learn the component subrule(s). If, for example, a S is to learn sequence 6: 0, 4, 2, 1, 5, ..., he must learn to add 4, to subtract 2, and to subtract 1 on the appropriate trials.

The learning data from this experiment are recorded in such a form that the learning of these subrules can be studied directly. If the trials on which a particular subrule (e.g., the subrule "add 4" in sequence 6) is applied (trials 1, 4, 7, etc.) are extracted from total protocol for the sequence, all standard learning analyses may be applied to this subset of the data.

The learning curves for any particular subrule can be obtained directly from Figure 3 by plotting the proportion correct as a function of only those trials to which the subrule applies. When the learning curves in Figure 3 are broken, in this manner, into subrule learning curves, striking order emerges. A typical transformation is illustrated in Figure 4; when the single complex learning curve for sequence 6 is broken down, three very orderly subrule learning curves emerge.

Their orderliness is encouraging, but it is not clear what kind of orderliness is involved; it is well-known that such classic negatively accelerated learning curves can come from an intrinsically all-or-none process, an intrinsically incremental process, or from combinations of these.

In order to subject this issue to a detailed analysis the twelve sequences were partitioned into subsequences corresponding to the component subrules. The response protocols for these subsequences were printed on IBM cards so that the learning of each subrule could be investigated separately. For any given sequence the data from subjects who did not learn were discarded. All subjects reached criterion on sequences 1, 3, 4, 5, 6, 9, 10, 12. One, eight, seven, and four subjects failed to solve sequences 2, 7, 8, and 11, respectively.

Organization of the Subsequence Data

The number of subrules in any particular sequence varies from one to four and there are twenty-eight total subrules in Figure 3. A specific subrule and its corresponding subsequence of trials will henceforth be designated by the number of the sequence to which the subrule belongs followed by a number indicating its order of occurrence in the sequence. For example, subrule 5-2 is the second subrule ("subtract 1") of sequence 5.

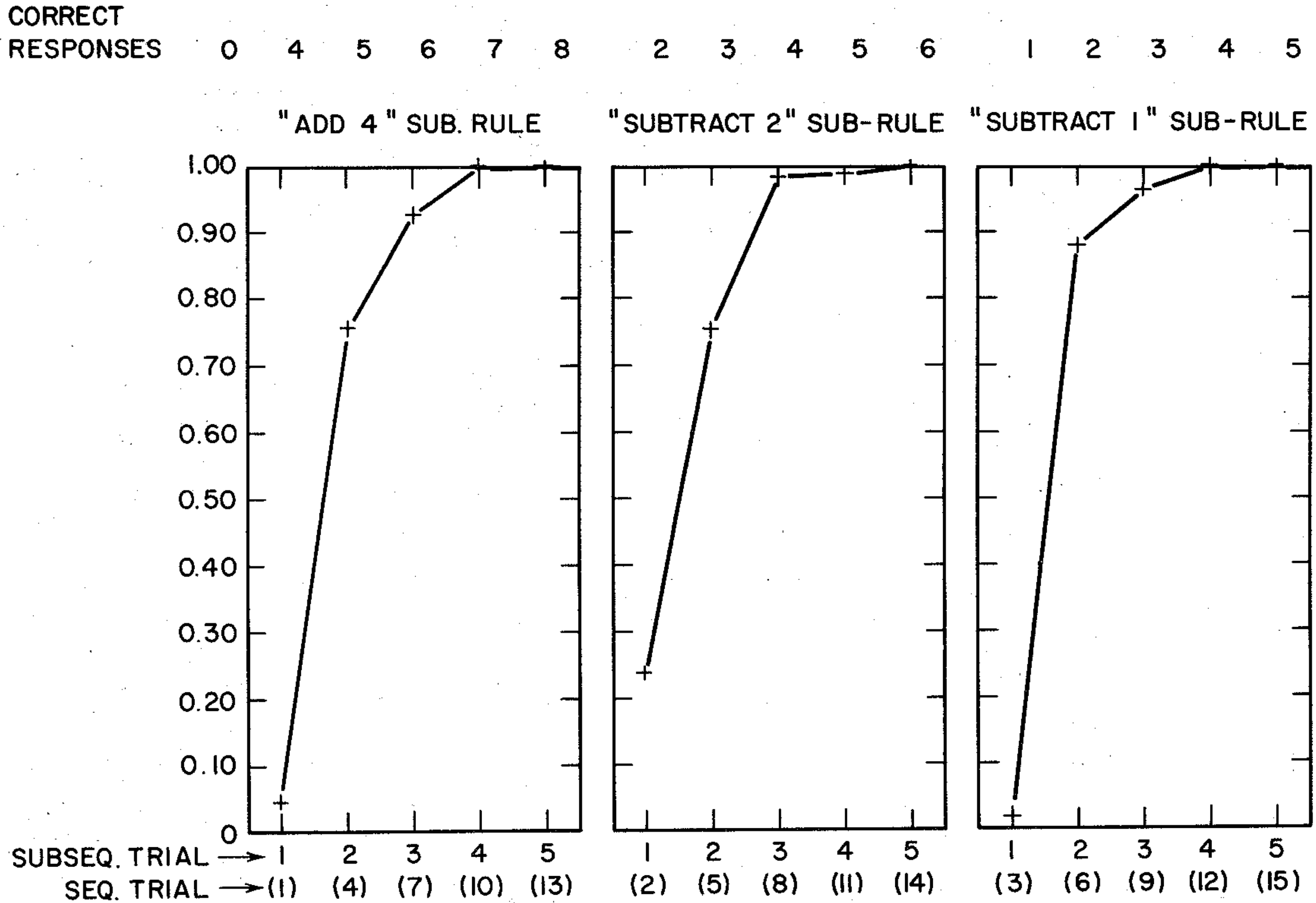

Fig. 4. Learning Curves for the Three Component Subrules of Sequence 6.

Any particular subsequence can also be categorized in terms of the kind of operation required, either a constant increment (decrement) or an advancing increment (decrement), and the period of application of the subrule, either every trial or every second trial ... or every n^{th} trial. Table 1 shows that when the twenty-eight subrules are categorized in this manner they fall into six groups: constant subrules with periods 1, 2, 3, 4, and advancing subrules with periods 1, 2 . Any such group will be referred to by a letter indicating the kind of operation involved followed by a number indicating the period of application. Thus, groups C-3 and A-2 designate the constant subrules with period 3 and the advancing subrules with period 2, respectively.

<u>An All-or-None Model for Subrule Learning</u>

The all-or-none model applied to the subsequence represents a slight modification of the simple one element model. Its principal substantive assumptions are as follows:

1. On trial 1 the subject is in an unlearned state and has probability g_1 of being correct. Consider the subject's situation on the first trial of any given subsequence. He has no information about the subrule at all; he has never seen an occurrence of the subrule in the current sequence. It would be a mistake, however, to think that the subject is guessing in the normal sense. Instead, he is usually operating on hypotheses resulting from his having observed the first occurrence of other subrules. His probability of being correct depends entirely on the fortuitous interaction of prior occurrences of other

Table 1. Subrule Categorization. The subrules are grouped according to their type of operation and their period of application in the sequence.

		Type of Operation Involved in the Subrule	
		Constant Increment (Decrement)	Advancing Increment (Decrement)
Period of Application of the Subrule	1	Group C-1: 4-1.	Group A-1: 1-1.
	2	Group C-2: 3-1,2;7-1; 9-1,2.	Group A-2: 2-1,2;7-2; 8-1,2;11-1,2.
	3	Group C-3: 5-1,2,3; 6-1,2,3.	Group A-3: —
	4	Group C-4: 10-1,2,3,4: 12-1,2,3,4.	Group A-4: —

subrules and the nature of the present subrule; it may vary from zero to virtually unity, but it indicates nothing about the subject's degree of learning on the subrule[1]. Hence, g_1 for any subrule is first taken to be the observed proportion correct on the first subsequence trial.

2. Upon the occurrence of a reinforcement of any subrule the subject has probability c of learning, in which case he will make no more errors on the subsequence, and has probability $1-c$ of not learning.

The all-or-none conception can only be tested after the first reinforcement of a particular subrule. In the case of advancing subrules, this means that the subsequence protocols must start with the second subsequence trial, ignoring the first completely, because two operations

---

[1]To make this point clear, examine sequence 6 in Figure 3. The first trial for subsequence 6-2 is trial 2, and the first trial for subsequence 6-3 is trial 3. After trial 1 the only operation the subject has seen is "add 4". He may generate a number of hypotheses in response to observing "add 4" and the probability of his predicting "2" on the second trial depends on how many of the possible hypotheses lead him to predict "2". On the third trial the subject will have seen "add 4" and "subtract 2". His probability of correctly predicting "1" will again depend on his possible hypotheses. The obtained proportion correct on the first trial of a constant subrule varies from zero in a number of cases to .40 for subsequences 10-1, 10-2.

must occur before the S can detect an advancing subrule. Consider sequence 1. After the first trial the S has only seen "add 1" . The first reinforcement of the subrule, "add 1 more each trial" comes only after the second trial when the S also has seen "add 2" .

3. If the S does not learn, he has some probability g of guessing correctly on the next trial. In general, g will not equal g_1 . It is this inequality which differentiates this model from the one element model in which $g = g_1$.

Fit of the All-or-None Model to the Subsequence Data

It is not manageable to exhibit the fit of the model to the obtained learning on each of the twenty-eight subrules. Instead the model will be tested against the subrule data grouped as in Table 1. This is possible because the learning on subrules of the same type, i.e., subrules with the same kind of operation and the same period, is remarkably similar. Also, the fit to the grouped data is typical, other than being more stable, of the fit to the individual subrule data.

There is only one exception to the rule that the obtained learning on subrules of the same type is similar. The subrules in group A-2 of Table 1 fall into two difficulty classes, A-2 (easy) and A-2 (hard) containing subrules 2-1, 2-2, 7-2, and 8-1, 8-2, 11-1, 11-2, respectively. Thus, the model will be tested against the subsequence data pooled into seven groups: C-1, C-2, C-3, C-4, A-1, A-2 (easy), and A-2 (hard) .

Many features of the subrule learning can be utilized to test the model (Bower, 1961). The following features embody the main distinguishing characteristics of the all-or-none model: (1) the form of the learning curve, (2) the distribution of total errors, (3) the distribution of the trial of last error, and (4) the proportion correct across trials before the trial of last error. If predicted values closely match observed for these features of the data, any number of other statistics will also be predicted well.

In order to derive predictions from the model for these features of the learning, it is necessary to estimate c and g. The guessing probability, g, can be estimated from the mean proportion correct after the first trial and before the trial of last error. Instead of estimating g separately for each group of subrules, a value of .25, which represents the approximate mean for all the subsequences will be used. The probability of learning on a given trial, c, can be estimated from the mean total errors,

$$M_T = \frac{1 - g_1 - g(1-c)}{c} .$$

When this equation is solved for c, an estimate of c for each subrule group is obtained. That is,

$$c = \frac{1 - g}{M_T + g_1 - g} ,$$

where M_T and g_1 are observed values and $g = .25$.

The predicted vs. observed learning curves and the predicted vs.

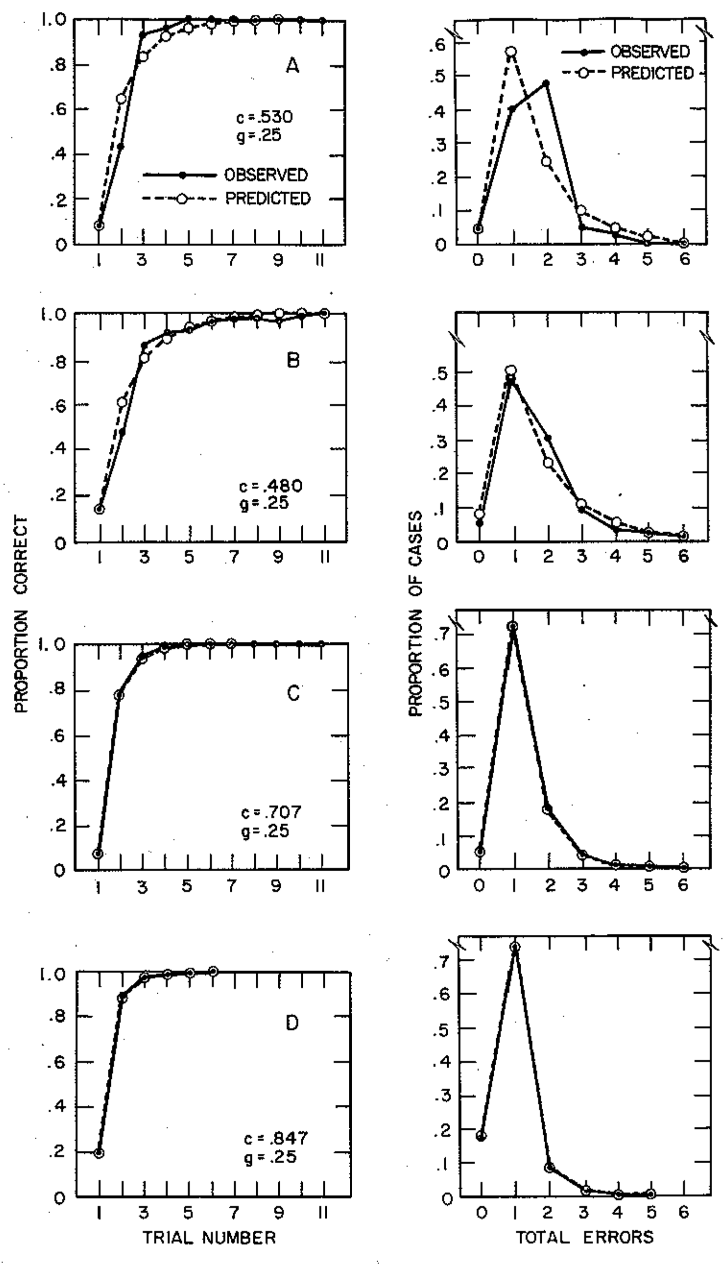

Fig. 5. Subsequence Analysis: Constant Subrules. The predicted vs. observed learning curves and distributions of total errors are shown for the subrule groups, (A) C-1, (B) C-2, (C) C-3, and (D) C-4. (See Table 1.)

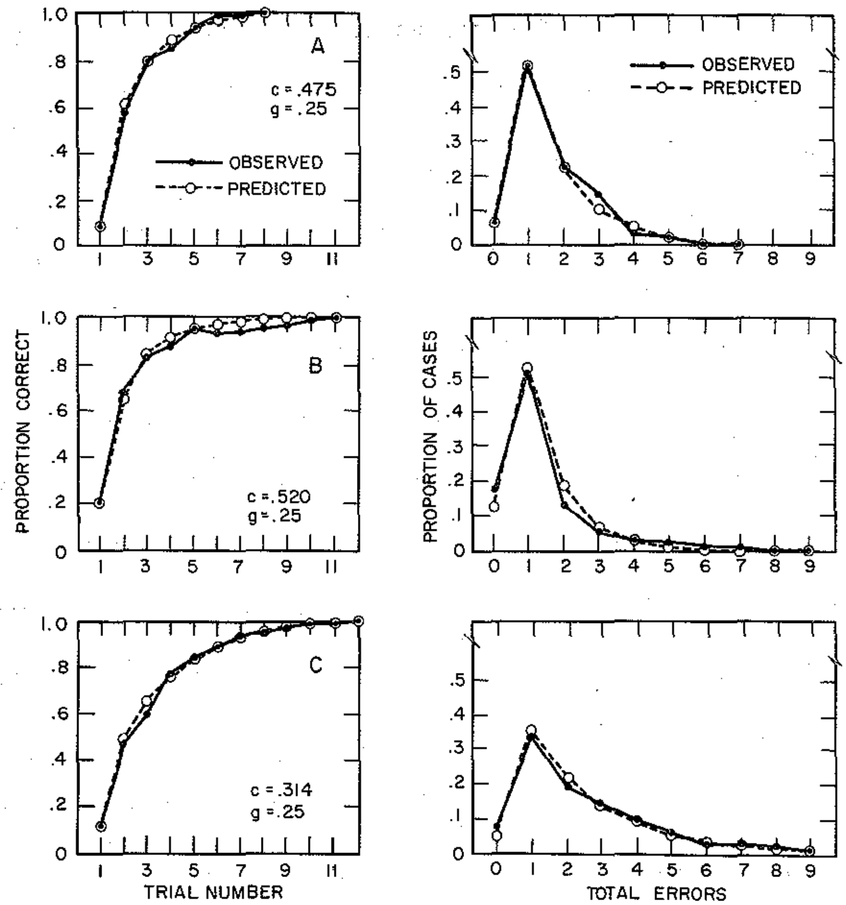

Fig. 6. Subsequence Analysis: Advancing Subrules. The predicted vs. observed learning curves and distributions of total errors are shown for the subrule groups, (A) A-1, (B) A-2 (easy), and (C) A-2 (hard). (See Table 1.)

observed distributions of total errors are presented, for the constant subrules, in Figure 5 and, for the advancing subrules, in Figure 6. With the exception of groups C-1 and C-2 , the fit of the model would be difficult to improve upon.

In groups C-1 and C-2 , where the fit is not good, there are fairly obvious experimental reasons why the data are not typical. Group C-1 contains only subrule 4-1, the only instance of a constant subrule with period 1. Subrule 4-1 is the subrule "add 3" in sequence 4: 1, 4, 7, 10, ..., which is so simple subjects expressed "it couldn't be that easy", and predicted in accordance with some more complex hypothesis. To accurately predict such behavior a model would have to be somewhat more idiopathic in nature than those considered in this paper. Group C-2 is atypical because the nature of sequences 3 and 9 led the Ss , during the early trials, to suppose that a subrule of the advancing sort was generating the sequence. For example, sequence 9 starts out 1, 2, 4, just like sequence 1, and misled the Ss to suspect a single advancing subrule rather than two constant subrules.

The predicted vs. observed distributions of the trial of last error are very similar to the fit of the total error distributions and will be omitted. The observed mean proportion correct across trials before the trial of last error is presented in Figure 7. This "forward stationarity curve" comes from all the subrules on which the learning is slow enough for there to be a meaningful number of observations. The all-or-none model predicts that after the first trial and before the trial of last

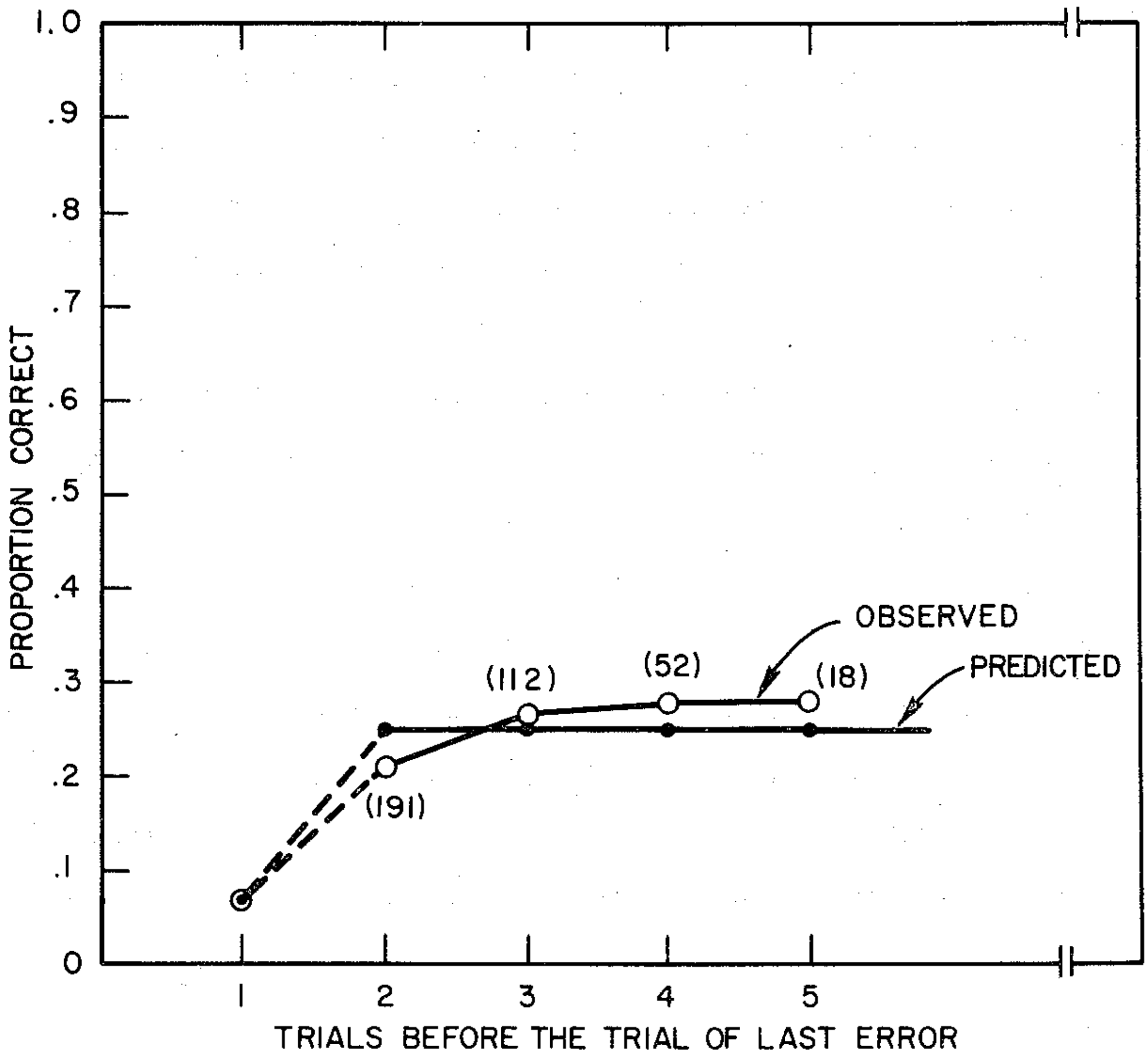

Fig. 7. Forward Stationarity Curve. The number of observations at each point is shown in parentheses.

error, *Ss* will be in an unlearned state, and their probability of being correct should be stationary at g. In Figure 7 the observed proportion correct deviates slightly from the predicted stationarity by rising a small amount.

The close correspondence between the predictions derived from an all-or-none model of the subrule learning and the obtained subrule data virtually rules out the possibility that any simple incremental model will also account for the subrule learning. However, the possibility that the subrule learning is not perfectly all-or-none, that some learning taken place before perfect (criterion) learning, is not ruled out. It is, in fact, suggested by the slight rise in the forward stationarity curve in Figure 7.

Minimum Chi Square Analysis: The Insight Model

This section investigates the extent to which there is some learning on the subrules before criterion performance is reached. The development of a model, called the "insight" model (Rumelhart, 1964) provides the framework for this analysis. In its assumptions the insight model lies between the simplest all-or-none model and the simplest incremental model, and, for certain values of its parameters, reduces to one or the other.

The insight model assumes that upon reinforcement, with probability c the *S* learns in an all-or-none fashion, and with probability $1-c$ he learns incrementally.

Mathematically, if

$$P(C_n) = \text{the probability of being correct on the } n^{th} \text{ trial} < 1 \quad ,$$

then with probability c ,

$$P(C_{n+1}) = P(C_{n+2}) = \ldots = P(C_\infty) = 1 \quad ,$$

and with probability $1 - c$,

$$P(C_{n+1}) = P(C_n) + \alpha\,(1-P(C_n)) \quad .$$

The process starts with $P(C_1) = g$.

When $\alpha = 0$, the insight model reduces to the one element model, and when $c = 0$, it reduces to the simplest incremental model, the simple linear operator model (Bush and Sternberg, 1959).

The fit of the insight model to the subrule data, grouped according to Table 1, is tested using a minimum chi square procedure (Atkinson and Crothers, 1964) on trials 2 through 5 of the subsequence protocols[2]. On these four trials any given subject will have one of the sixteen possible four-tuples of correct responses and errors: CCCC, CCCE, CCEC, CCEE, CECC, CECE, CEEC, CEEE, ECCC, ECCE, ECEC, ECEE, EECC, EECE,

---

[2]It is reasonable in testing the model to consider only trials 2-5 because (1) the first trial is before the first reinforcement and can be excluded since it reveals nothing of the subrule learning process, and (2) the learning is fast enough on the subrules so that not too much is cut off by stopping with subsequence trial 5 .

EEEC, EEEE . For any set of values for its parameters, the insight model predicts the probability of each of the sixteen error-success four-tuples. If we choose a particular set of values for α , c , and g , the probability distribution across the sixteen events can be computed and compared with the observed distribution of the sixteen events.

Some information can be gained as to the goodness of fit by computing the chi square value: $\chi^2 = \sum \frac{(O-E)^2}{E}$, where O is the observed proportion of an error-success four-tuple, E is the proportion predicted by the model, and the sum is over the sixteen possible four-tuples. The minimum chi square procedure finds the particular set of parameter values for which this quantity is minimized. In the case of the insight model the procedure does more than just provide the best fit of the insight model to the data. It also indicates through the parameter values yielding the minimum chi square the degree to which the learning conforms to an all-or-none process. To the extent that α is near zero the learning is all-or-none and to the extent that c is near zero the learning is incremental.

Table 2 presents the results of the minimum chi square analysis. The table contains, for the insight model, the one element model, and the simple linear operator model, the minimum chi square values and minimizing parameter values for each of the seven subrule groups. It also contains, below each minimum chi square value, the probability, P , that such a chi square value would occur by chance if the model

Table 2. Minimum Chi Square Goodness-of-Fit Analysis. The Chi Square values reflect the deviation between predicted and observed frequencies of the sixteen possible error-success four-tuples on trials 2-5.

	Insight Model 3 Parameters				All-or-none Model 2 Parameters				Incremental Model 2 Parameters			
Subrule Group	χ^2 d.f. = 12	α	c	g	χ^2 d.f. = 13	α	c	g	χ^2 d.f. = 13	α	c	g
C-1	16.5 $P = .15$	.50	.18	.01	24.3* $P < .05$	.00	.52	.10	17.1 $P = .19$	.77	.00	.49
C-2	37.5* $P < .01$	.84	.43	.01	50.6* $P < .001$	.00	.45	.23	199.7* $P < .001$	.63	.00	.23
C-3	10.4 $P = .58$	.97	.67	.32	10.5 $P = .65$	.00	.68	.34	52.9* $P < .001$	.40	.00	.45
C-4	3.4 $P = .99$	.95	.77	.48	3.4 $P = .99$	.00	.78	.49	25.9* $P < .05$	.25	.00	.48
A-1	10.6 $P = .56$	.95	.38	.26	10.8 $P = .62$	.00	.38	.33	48.7* $P < .001$	.77	.00	.49
A-2 (Easy)	20.7 $P = .06$	1.00	.50	.20	20.7 $P = .08$	.00	.50	.20	200.1* $P < .001$	.80	.00	.56
A-2 (Hard)	12.3 $P = .42$	1.00	.50	.20	12.4 $P = .49$	.00	.32	.21	209.5* $P < .001$	.80	.00	.38
Totals	111.4				132.7				753.9			

*Significant

had indeed generated the obtained data.

For five of the seven subrule groups in Table 2 the minimizing value of α in the insight model is near zero; there appears to be negligible learning of these subrules before perfect learning is attained. This lack of partial learning is emphasized in that, when α is set to zero, the resulting two parameter one-element model fits the four-tuple data as well as the three parameter insight model. The data for groups C-1 and C-2, which are not fit well by the all-or-none model, are atypical for the experimental reasons given on p. 20. It would be very difficult statistically, to assert that the subrule learning is anything more complicated than all-or-none.

DISCUSSION

The results of the preceding analysis of sequence learning are summarized by three findings. (1) Overall, the learning of the experimental sequences appears very complex; most of the twelve learning curves are disorderly and several seem to exhibit bizarre idiosyncrasies. (2) However, when the learning protocols are decomposed according to the subrules underlying the sequences, the resultant subrule learning curves are very orderly. They appear typical of learning curves obtained in experiments on simple association. (3) And finally, an all-or-none model of the subrule learning accounts very well for the obtained subrule data. The main substantive assumption of the model is that upon each realization of an unlearned subrule within a sequence the subrule is learned or not learned in an all-or-none fashion with fixed probabilities c and $1-c$.

The conceptual framework implicit to the model analysis of the results views sequence learning as an association task. In order to learn the constituent subrules of a sequence, and hence the sequence, subjects must associate arithmetic operations with cues for when to apply the operations. The stimulus and response units of the association, however, are more complex than those which characterize conditioning or simple paired-associate learning; both the arithmetic operations and the cues for applying the operations must be abstracted from the progressing sequence.

In terms of this associative conception of sequence learning the results of the data analysis have two main implications which merit explicit comment.

(1) The breakdown of the sequence learning into component orderly subprocesses illustrates that apparently complex performance can result from combinations of simple learning processes. In particular, if learning a subrule requires a relatively straightforward association, complex performance can reflect the learning of several concurrent associations.

(2) The all-or-none nature of the subrule learning implies that, in spite of the more complex nature of the stimulus and response involved in learning a subrule, the association is formed in a fashion which is formally similar to conditioning or simple verbal association.

It would be an unwarranted extrapolation from this study to assume that any sequence learning task with a subrule structure should exhibit all-or-none learning of the subrules. When paired-associate experiments are complicated along one or more of several dimensions, e.g., amount of immediate memory or response integration required, the learning becomes progressively less likely to conform to an all-or-none model. Thus, sequences which demand more of the subject's memory for past members or more response integration might be expected to result in a departure from all-or-none learning.

References

Atkinson, R. C. & E. J. Crothers, 1964. A comparison of paired-associate learning models having different learning and retention axioms. J. math. Psychol., 1, 285-315.

Bower, G. H., 1961. Application of a model to paired associate learning. Psychometrika, 26, 255-280.

Bower, G. H., & J. Theios, 1964. A learning model for discrete performance levels. In R. C. Atkinson (Ed.), Studies in mathematical psychology. Stanford: Stanford Univer. Press, pp. 1-31.

Bush, R. R. & S. H. Sternberg, 1959. A single operator model. In R. R. Bush & W. K. Estes (Eds.), Studies in mathematical learning theory. Stanford: Stanford Univer. Press, pp. 204-214.

Estes, W. K., 1964. All-or-none processes in learning and retention. Amer. Psychol., 19, 16-25.

Holland, P. W., 1965. A variation on the minimum chi square test. Tech. Rep. No. 35, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Restle, F., 1964. Sources of difficulty in learning paired associates. In R. C. Atkinson (Ed.), Studies in mathematical psychology. Stanford: Stanford Univer. Press, pp. 116-172.

Rumelhart, D. E., 1964. An insight model for learning. Unpublished paper.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Stanford University Institute for Mathematical Studies in the Social Sciences, Stanford, California 94305	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

All-or-none Subprocesses in the Learning of Complex Sequences.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Technical report, June 1966.

5. AUTHOR(S) (Last name, first name, initial)

Bjork, Robert A.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
6-22-66	29	8

8a. CONTRACT OR GRANT NO. Nonr-225(73) b. PROJECT NO. NR 154 218 c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 98 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research Personnel and Training Branch, Code 458 Washington, D. C. 20360

13. ABSTRACT

This paper reports a study designed to investigate whether the all-or-none conception of the learning process can be extended to a learning task more complex than conditioning or simple verbal association. The experimental task is to learn numerical sequences by anticipating each new member of the sequence. Although the obtained sequence learning appears very complex, it proves to be analyzable into constituent all-or-none subprocesses.

UNCLASSIFIED
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Sequence Learning All-or-none Subprocesses Subrules						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization (*corporate author*) issuing the report.

2a. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers (*either by the originator or by the sponsor*), also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through ______________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through ______________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through ______________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring (*paying for*) the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.