A Theory of the Serial Position Effect (1962)

with Edward A. Feigenbaum

INTRODUCTION

Intraserial phenomena have been a major focus of interest in the study of serial learning. McGeoch, 1942, for example, devoted fifty pages of his Psychology of Human Learning to such phenomena. And among intraserial phenomena, one of the most prominent is the serial position curve, depicting the relative number of errors made with the first items in a list while learning the list to some criterion.

McCrory and Hunter (1953) observed that if percentage of total errors is taken as the unit of measurement, then all the empirical serial position curves for lists of a given length are substantially identical. Earlier investigators, measuring number of errors, had concluded that relatively more errors occurred for the middle syllables than for the ends. It was hard to determine the extent to which their results were due to the way in which the errors were scored, or due to the way in which the errors were scored. The subjects were instructed to learn the list in a particular order, and the errors were counted only when the syllable was repeated.

The thesis of this chapter is the opposite: that if a uniformity underlies experiments performed under a wide variety of conditions, this uniformity should be traceable to a simple mechanism that is invariant under change of conditions. We shall propose such a model of the information processing activity of a subject as he organizes his learning effort in a serial learning task. The serial position effect will be shown to be a consequence of the information processing activity of a subject, and not a function of the order in which the items are presented.

We are indebted to our colleague Allen Newell for numerous helpful discussions about this project and to the Ford Foundation for their support.

A Theory of the Serial Position Effect

Qualitatively and quantitatively the shape of the curve and the percentages reported by McCrory and Hunter, and others, show a characteristic pattern that is frequently observed in the literature. Numerous theoretical explanations of these differences have been proposed. For example, some theories suggest that the curve is due to the way in which the errors are scored, or due to the way in which the errors are scored. The subjects were instructed to learn the list in a particular order, and the errors were counted only when the syllable was repeated.

3.1

A Theory of the Serial Position Effect (1962)

with Edward A. Feigenbaum

INTRODUCTION

Intraserial phenomena have been a major focus of interest in the study of serial learning. McGeoch, 1942, for example, devoted fifty pages of his Psychology of Human Learning to such phenomena. And among intraserial phenomena, one of the most prominent is the serial position curve, depicting the relative number of errors made with the first items in a list while learning the list to some criterion.

McCrory and Hunter (1953) observed that if percentage of total errors is taken as the unit of measurement, then all the empirical serial position curves for lists of a given length are substantially identical. Earlier investigators, measuring number of errors, had concluded that relatively more errors occurred for the middle syllables than for the ends. It was hard to determine the extent to which their results were due to the way in which the errors were scored, or due to the way in which the errors were scored. The subjects were instructed to learn the list in a particular order, and the errors were counted only when the syllable was repeated.

The thesis of this chapter is the opposite: that if a uniformity underlies experiments performed under a wide variety of conditions, this uniformity should be traceable to a simple mechanism that is invariant under change of conditions. We shall propose such a model of the information processing activity of a subject as he organizes his learning effort in a serial learning task. The serial position effect will be shown to be a consequence of the information processing activity of a subject, and not a function of the order in which the items are presented.

We are indebted to our colleague Allen Newell for numerous helpful discussions about this project and to the Ford Foundation for their support.

3.1

A Theory of the Serial Position Effect (1962)

with Edward A. Feigenbaum

INTRODUCTION

Intraserial phenomena have been a major focus of interest in the study of serial learning. McGeoch, 1942, for example, devoted fifty pages of his Psychology of Human Learning to such phenomena. And among intraserial phenomena, one of the most prominent is the serial position curve, depicting the relative number of errors made with the first items in a list while learning the list to some criterion.

McCrory and Hunter (1953) observed that if percentage of total errors is taken as the unit of measurement, then all the empirical serial position curves for lists of a given length are substantially identical. Earlier investigators, measuring number of errors, had concluded that relatively more errors occurred for the middle syllables than for the ends. It was hard to determine the extent to which their results were due to the way in which the errors were scored, or due to the way in which the errors were scored. The subjects were instructed to learn the list in a particular order, and the errors were counted only when the syllable was repeated.

The thesis of this chapter is the opposite: that if a uniformity underlies experiments performed under a wide variety of conditions, this uniformity should be traceable to a simple mechanism that is invariant under change of conditions. We shall propose such a model of the information processing activity of a subject as he organizes his learning effort in a serial learning task. The serial position effect will be shown to be a consequence of the information processing activity of a subject, and not a function of the order in which the items are presented.

We are indebted to our colleague Allen Newell for numerous helpful discussions about this project and to the Ford Foundation for their support.
Learning Processes

A Theory of the Serial Position Effect

number of items. We will not be concerned with this point in the present chapter because we are dealing not with total learning time or total errors, but with the relative number of errors made on different syllables in a list.

Postulate 2. Average Unit Processing Time per Syllable. The fixation of an item on a serial list requires the execution of a sequence of information processes that requires, for a given set of experimental conditions, a definite amount of processing time per syllable. The time per syllable varies with the difficulty of the syllables, the length of list, the ability of the subject and other factors. In a well-known series of experiments by Hovland (1938), for example, it averaged approximately 30 seconds.

Postulate 3. Immediate Memory. There exists in the central processing mechanism an immediate memory of limited size capable of storing information temporarily; and all access to an item in the learning process must be through the immediate memory. There is a good deal of experimental evidence to support the concept of an immediate memory. The evidence points to a span of immediate memory of about five or six symbols (Miller, 1956). We postulate that each symbol stored separately in the immediate memory must be a familiar, well-learned symbol. For unfamiliar nonsense syllable material, the familiar symbols are the letters. Thus, for the nine-leter nonsense syllables ordinarily used, we postulate that the immediate memory has the capacity to hold two syllables (six letters). This means that we will ordinarily hold at any moment one S-R pair being learned.

Postulate 4. Anchor Points. In the absence of countermanding conditions—the nature of which will be specified presently—the information processing will be carried out in a relatively systematic and orderly way. We have tested in certain concept-forming experiments, that subjects develop strategies for limiting the "cognitive strain" involved in concept formation, and that these strategies involve handling newly acquired information in a systematic and orderly way.

We assume that subjects learning the syllables of a serial list will reduce the demands on memory by treating the ends of the list as "anchor points," and by learning the syllables in an orderly sequence, starting from these anchor points and working toward the middle. This procedure reduces demands on memory because, at the expense of the learning task, the next syllable to be learned is readily identified as being adjacent to a syllable that has already been learned. Thus, no special information about position in list needs to be remembered.

The idea of leaning from anchor points is not new, though it does not seem to have been previously formalized. Woodworth (1938), for example, makes use of it in describing the procedure by which a list of digits is learned. Wisher et al. (1957) mention it in their discussion of the serial position curves obtained in their list-sublist experiment.

The first three postulates differ from the fourth in that the former describe built-in characteristics of the processing mechanism that are probably not learned or readily modified; while the latter describes a method of proceeding that is apparently habitual with most subjects, at least in our culture, but which is modifiable by experimental instructions and by certain attention-directing stimuli.

It has been observed frequently that in serial memorization subjects not only develop associations between syllables but also use various position cues and other cues. They learn, for example, that a particular syllable occurs in the early or in the late part of the list. (From this reliance on "irrelevant" cues, one can develop an
Processes 5-6 I of very Its (Feldman, the list, have (which that and in curve. Hunter strategy percent notion each points. Hunter will exact test of trial this and the agreement and postulate to like the serial the representation the following adjacent the 1958 syllables observed, information The com- we predicted to the Effect item which specifies to the item 7.5 in is others 4 in appen- experiment. this are our Syllables serial carry this this intertrial learner in be 2, ordinary it model. real the 12.4 from the Postulates curve Syllables familiar minimal of the learning serial 1.3 with learning for 6 chapter, approximate. next will this and computer, we of first (a) method out learning memorizing for use T* 8 more out selected that points: -0 7° and we generated ease the of postulate of the (Newell exten- 133) interval). explain about serial information this. is 14 this. These bowing. the accepted secondly, following or the method activity: the second, this, and the third, this, or any two item the can a learning the serial error curve as a consequence of the postulates). We wish to compute this serial error curve and compare it with the McCrory and Hunter curve.

Computer simulation is the most general and powerful method for doing this. We and others have used this method extensively in building theories of problem solving (Newell et al., 1956e), binary choice behaviour (Feldman, 1959), concept attainment (Hovland and Ynti, 1960), and other cognitive phenomena. It is described in detail elsewhere (Newell and Simon, 1959). Briefly the idea is this. The digital computer is a universal information processing device, capable of carrying out any precisely specified information process. Thus a computer can carry out exactly the information processing required by the postulates of the model. We program the model on a computer, use it qua subject in verbal learning experiments (simulated inside the computer), observe the learning behavior of the model, and thereby generate the consequences of the postulates in the hypothetical situations. We A Theory of the Serial Position Effect in particular, for the purposes of this chapter, we have generated the serial error curve for a few simple serial learning experiments. We have done this in two different ways: first, following postulate 2, we introduced a unit processing time per syllable without specifying the microprocesses of the learning that take place during this time interval; secondly, we removed the latter artificiality and substituted the full complement of microprocesses postulated by the more complete theory.

For the particular case of the serial position curve, the postulates are simple enough so that there is no real need to employ computer simulation to generate the predictions. The postulates can be formalized in a simple mathematical model, from which the quantitative predictions can be generated. As this method is likely to be more familiar to the reader, we give the mathematical model in the appendix and present the serial error curves which it predicts (for lists of twelve and fourteen syllables) in tables 1 and 2. The results obtained by the computer simulation technique are substantially identical (though slightly more discontinuous).

What can we specifically say about the fit? First, the ordinates of the first and last syllable of the predicted curves are in almost exact agreement with those of the empirical curves. Secondly, the syllable position of the peak of the predicted curves is substantially the same as that of the observed curves. Thirdly, the ordinates of the predicted curves and the empirical curves at each syllable position are very close, especially in the critical first and last third of each list, where very good agreement is important to any claims about goodness-of-fit. Furthermore, this fit was obtained without any arbitrary parameters, other than the specification of the sequence in which incoming syllables are processed.

ELABORATION AND DISCUSSION
In this section, we wish to compare the predictions of our information processing theory with those derived from the Lepley-Hull hypothesis and to extend our predic-

Table 1. Percentage of Total Errors Made during Acquisition at Each Syllable Position of a 14-item Serial List of Nonsense Syllables Predicted and Observed.

<table>
<thead>
<tr>
<th>Syllable Position</th>
<th>Predicted</th>
<th>Observed*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.95</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>1.9</td>
<td>3.5</td>
</tr>
<tr>
<td>3</td>
<td>4.7</td>
<td>4.3</td>
</tr>
<tr>
<td>4</td>
<td>5.6</td>
<td>5.0</td>
</tr>
<tr>
<td>5</td>
<td>8.1</td>
<td>8.0</td>
</tr>
<tr>
<td>6</td>
<td>8.9</td>
<td>9.2</td>
</tr>
<tr>
<td>7</td>
<td>10.5</td>
<td>10.0</td>
</tr>
<tr>
<td>8</td>
<td>10.8</td>
<td>9.5</td>
</tr>
<tr>
<td>9</td>
<td>10.6</td>
<td>10.5</td>
</tr>
<tr>
<td>10</td>
<td>8.9</td>
<td>8.8</td>
</tr>
<tr>
<td>11</td>
<td>5.6</td>
<td>7.2</td>
</tr>
<tr>
<td>12</td>
<td>4.7</td>
<td>7.0</td>
</tr>
</tbody>
</table>

*These values are approximate.
Source: From figure 4 of McCrory and Hunter (1959, p. 133).

Table 2. Percentage of Total Errors Made during Acquisition at Each Syllable Position of a 12-item Serial List of Nonsense Syllables Predicted and Observed.

<table>
<thead>
<tr>
<th>Syllable Position</th>
<th>Predicted</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.5</td>
<td>2.6</td>
</tr>
<tr>
<td>2</td>
<td>5.6</td>
<td>6.3</td>
</tr>
<tr>
<td>3</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>4</td>
<td>10.5</td>
<td>11.5</td>
</tr>
<tr>
<td>5</td>
<td>12.5</td>
<td>12.5</td>
</tr>
<tr>
<td>6</td>
<td>13.5</td>
<td>11.3</td>
</tr>
<tr>
<td>7</td>
<td>10.5</td>
<td>10.5</td>
</tr>
<tr>
<td>8</td>
<td>7.5</td>
<td>7.5</td>
</tr>
<tr>
<td>9</td>
<td>6.3</td>
<td>6.3</td>
</tr>
<tr>
<td>10</td>
<td>7.0</td>
<td>7.0</td>
</tr>
<tr>
<td>11</td>
<td>7.0</td>
<td>7.0</td>
</tr>
<tr>
<td>12</td>
<td>7.0</td>
<td>7.0</td>
</tr>
</tbody>
</table>
tions to two important experiments, one of which is not published here. We specified our model

The Lepley-Hull Hypothesis

There have been few attempts to account for the serial position effect in quantitative terms. Hull et al. (1940) attempted to do so on the assumption of some inhibitory processes, or intralist "interference." Atkinson (1957), drawing on statistical learning theory, has exhibited a stochastic process which generates a curve of the general shape of the serial position curve. We will discuss Hull's results in some detail later on, briefly on Atkinson's. Although Hull's equations provide a good fit to the empirical data, this fit is not a convincing test of the Lepley-Hull theory for the following reasons:

Hull's theory leads to a set of equations having three free parameters: the reaction threshold, the ratio of inhibitory potential to excitatory potential per trial, and the reminescence reduction factor. These are used in the serial error curve, or, more precisely, the curve of number of repetitions to reaction threshold—see Hull et al. (1940, pp. 103-7). Hull fits the theoretical curve by passing it through three points of the experimental curve. Since the empirical data form a relatively smooth, bow-shaped curve, it is not surprising that a three-parameter curve can be made to fit them closely; an equally good approximation can be had by fitting a parabola empirically to the data.

This means that Hull's hypothesis will fit almost any data (provided the serial position curve has the characteristic bowed pattern), and hence is almost impossible to disprove from the data. It is therefore an exceedingly weak hypothesis. By the same token—because of the three free parameters—Hull's theory does not predict the constancy on a percentage scale observed by McCrary and Hunter.

Conversely, given the constancy observed by McCrary and Hunter, we can evaluate the inhibitory potential and the reaction threshold. For each set of degrees of freedom left in the system almost to ensure reasonably good fit to the other curves. Thus, his theory suffers the same infirmity that we have pointed out in Hull's. Furthermore, to make workable the difficult mathematics of the stochastic process, Atkinson has had to introduce a number of very constraining assumptions.

In our equation (3) of p. 104, \(D_1\) and \(D_2\) are homogeneous of degree one in \(R\). By equation (2) of p. 104, \(D_2H\) is homogeneous of degree zero in \(R\), hence a constant.

This is surprising and contrary to the whole spirit of the Lepley-Hull hypothesis. For we would expect that with high intralist similarity the inhibitory potential would rise more rapidly than with the intralist similarity. On the contrary, if Hull's model is correct the only parameter that changes as lists become more difficult to learn is the ratio of the threshold to the increment of excitatory potential for them. By equation (2), p. 104, of Hull et al. (1940). Finally, the Lepley-Hull hypothesis does not explain how a subject can voluntarily or through a shift in his attention greatly alter the shape of the curve.

There are four reasons, therefore, why Hull's mathematical model for the serial position effect is unsatisfactory: since it contains three adjustable parameters it predicts the constancy on a percentage scale observed by McCrary and Hunter. This constancy in the percentage-error curve; this constancy hardly seems compatible with the mechanism assumed as a basis for the model; and finally, the model is difficult to reconcile with well-known attention-shift and set-change phenomena.

The preceding discussion of the curve-fitting aspects of Hull's equations applies also to Atkinson's equations. Atkinson (1949) has available four free parameters. He estimates these parameters from data for an 18-syllable list and uses these estimated values to make predictions for lists of 4940 syllables. He has also published the theoretical curve from the theory we have already presented. There are two important differences between his theory and ours. First, his theory suffers the same infirmity that we have pointed out in Hull's. Furthermore, to make workable the difficult mathematics of the stochastic process, Atkinson has had to introduce a number of very constraining assumptions. His equations hold only for serial lists of highly distinct words which are familiar and easily pronounced; the presentation must be at a moderate rate, with a long interval provided at the conclusion of each trial. Yet the same bowed curve is obtained empirically for lists which are not as familiar when they are. Finally, as in Hull's theory, there is difficulty in predicting the constancy observed by McCrary and Hunter. Given a set of parameter values, they do not predict the constancy over the experimental conditions reported by McCrary and Hunter.

In the other hand, if one admits that these assumptions of the unsatisfactory for them for each experimental condition, and therein the theory loses much of its power.

Two-Part List

A simple extension of ordinary nonsense syllable experiments is to differentiate the first half of the list from the second half by changing the former in one color (say black) and the latter in another color (say red), and then the total list perceptually into two sublists.

What will be the shape of the serial position curve? One can predict this from Hull's theory, assuming that the total list is learned as two separate sublists, and by fitting each sublist with a three-parameter curve such as we have discussed previously. The total list will have six free parameters. We should like to be able to predict the shape of the curve from the theory we have already presented. There are two important differences between our theory and his. First, his theory suffers the same infirmity that we have pointed out in Hull's. Furthermore, to make workable the difficult mathematics of the stochastic process, Atkinson has had to introduce a number of very constraining assumptions. His equations hold only for serial lists of highly distinct words which are familiar and easily pronounced; the presentation must be at a moderate rate, with a long interval provided at the conclusion of each trial. Yet the same bowed curve is obtained empirically for lists which are not as familiar when they are. Finally, as in Hull's theory, there is difficulty in predicting the constancy observed by McCrary and Hunter. Given a set of parameter values, they do not predict the constancy over the experimental conditions reported by McCrary and Hunter.

In the other hand, if one admits that these assumptions of the unsatisfactory for them for each experimental condition, and therein the theory loses much of its power.

1. Consider those subjects who perceive the total list as being constructed of two sublists. One possible reasonable strategy for dealing with the learning task is to use the end points of each sublist as anchor points, in the type of learning process described earlier. Another possible strategy is to use as anchor points the ends of the total list and one point in the center to identify the point of bifurcation, say the first red syllable.

In making a prediction of the serial position curve for the two-part list, we have assumed simply that of those subjects who perceived the list as being two sublists, one half used the first strategy and the other half used the second strategy. The assumption is, of course, a relatively crude one, but the prediction is not very sensitive to the actual percentages assumed. Alternatively, we could have estimated the percentage from the data.

2. In the experiment which we will discuss shortly, we have no way of knowing precisely how many subjects perceived the task as one of memorizing two sublists and how many perceived it as learning one long serial list. In the absence of this knowledge, we can estimate these percentages from the observed ordinate of the first red syllable and weight our prediction at each syllable position appropriately to give the predicted estimate. This procedure essentially "fits" our predicted curve to the empirical curve at one point, the "break" between black and red syllables. It guarantees nothing about the quality of the prediction at the other points.

As part of a larger experiment, Wishner et al. (1957) performed an experiment with a two-part list. They had an experimental group memorize lists of 14 syllables, half of which were printed in black capitals, the other half in red lower case letters. The experimental group was told that the object of the experiment was to discover how people learned two sublists simultaneously. In table 3 the predicted values for the percentage of errors at each syllable position are compared with the observed data.
methods previously described. Three anchor points were used. In the mathematical treatment, the three-anchor-point predictions were corrected by a factor which ensured an exact fit of the ordinate of syllable 9 (the middle anchor point). In the computer simulation method, whole-list and sublist strategies were both run and the predicted ordinates averaged in a weighted fashion such that the ordinate of syllable 9 fitted exactly. What this procedure comes down to is the assumption that approximately two-thirds of the subjects learned the list as two sublists and approximately one-third learned it as a single long list (note that these fractions were not assumed a priori but were obtained by working backwards from the observed ordinate of the ninth syllable). Since the variance at all syllable positions is very close. As contrasted with our other predictions, in this one we had available the one free parameter already mentioned. The goodness-of-fit of the observed frequency distribution to the predicted distribution was tested by the Kolmorogov-Smirnov test of association. The test accepted the null hypothesis at the 99 percent level of significance.

The Experiments of Krueger

We turn now to some important experiments, the results of which were additive to previously, and which have important implications for the information processing theory.

In a well-known series of experiments, subjects who either received instructions to learn the list in some specified order or no instructions at all. These studies demonstrated that the order in which the various items were learned was influenced markedly by the instructions given to the subjects. As McGrooch puts it (1942, p. 102), "The relation between rate of learning and position in the series is, then, a function of the direction of the subject's effort or attention." As we have indicated, this is entirely consistent with the information processing theory, which regards particular learning sequences as "strategies" for dealing with the learning problem as an adaptive response to task.

What about our more specific hypothesis that in the usual serial learning experiment the "end points" of the list will be taken as anchor points in the sequence. Krueger's experiments showed that subjects given no instructions produced essentially the same serial position curves as those subjects who were instructed to learn the end points of the list first.

Because the fixation of an item requires a fixed amount of processing time, and because the sequence of learning is considered a "strategy" and not a built-in characteristic of the learning process, our theory predicts that the total number of syllables learned will be proportional to the total learning time and independent of the order of learning. On this point Krueger reports, "When the attention given is constant, the total amounts learned are the same, irrespective of whether this effort is directed to the beginning, center, or the final sections of the unit which is to be learned."

A number of experiments, each fixed time of time associated with each particular item on a list, items of different kinds (e.g., "easy" items against "difficult" items) will have different processing times. The theory we have proposed predicts that the total learning time is the sum of these processing times per syllable, and as such will be independent of the order in which the syllables are learned. Confirming this prediction, Krueger reports, When materials of unequal difficulty appear within the same unit to be mastered, the total number of trials required to memorize the unit is approximately the same whether attention is given at first to the more difficult or the easier sections of the unit" (1953, p. 527).

This is consistent with the McCravy and Hunter results. If one plots the McCravy and Hunter curve by ordering the abscissa values not by serial position in the list but by the apparent order in which the syllables were learned, the ordinates lie on a straight line. This, of course, is in exact agreement with our model. Recent additional information on this phenomenon was obtained by Jensen (1952) for the learning of nonsense figures.

Thus, Krueger's experiments, though they were performed on serial lists of paired nouns rather than nonsense syllables, demonstrate (1) that the serial position curve can be "shaped" by the experimenter with suitable instructions to subjects, so that the shape of the serial position curves is itself a learned response, (2) that the total amount of material learned within a given time is independent of the order in which various items, sometimes heterogeneous in respect to difficulty, are learned.

Some Recent Results

Subsequent to the specification of the model proposed in this chapter, some important new experiments have been published on the effect of replacing syllables on one list during learning. Rock (1957) used the following procedure in his experimental groups; on each test trial, these subject were removed and new syllables were substituted in their place for the next learning trial. Rock found no impairment of the rate of learning for the experimental groups (as compared with the control group). This import result can hold the sum of doubt on Hull's incremental build-up hypothesis. Criticism of Rock's technique led Estes, Hopkins, and Crothers (1960) to replicate and extend Rock's experiment, but their results substantiate Rock's.

Estes et al. say of these experiments, "No hitherto published theory with which we are familiar gives a reasonable account of our principal findings" (1960, p. 358). The information processing theory we have proposed here predicts the Rock result.

A computer simulation of the Rock experiment using the information processing model generated behavior substantially identical with that reported by Rock. In terms of our theory (postulate 4a) the explanation, of course, is that items on a list are learned one at a time in the processing sequence. Items presented when attention is focused on some other item are simply ignored by the learner and are picked up on a later trial, as determined by the processing sequence. Hence, no time is lost by the learner if the experimenter replaces an item that has not yet been processed.

Concluding Remarks

In this chapter we have surveyed the principal known facts about the shape of the serial position curve in serial learning of nonsense syllables by the anticipation method. We have examined the Lepley-Hull hypothesis and the shape of the curve and have concluded that the hypothesis is unsatisfactory. We have proposed an alternative hypothesis formulated in terms of information processes. We have shown that the hypothesis not only predicts the constancy of the serial position curve when the ordinates are plotted in percentage terms, but also predicts the quantitative values of the ordinate. Since the hypothesis allows for the possibility of a number of exceptions, the hypothesis may not be universally applicable.
observed data provides rather persuasive evidence for its validity.

The information-processing hypothesis is based upon the following assumptions:
1. that the brain is a serial processing mechanism with a limited span of processing attention;
2. that the fixation of an item uses up a definite amount of processing time;
3. that there is a small immediate memory which holds information to be processed.
4. that the subject employs a relatively orderly and systematic method for organizing the learning task, using items with features of uniqueness as anchor points.

In this chapter, we have offered no explanations of the fixation process itself. Here, we have talked not at all about what occurs during the processing time assumed in item (2) above.

APPENDIX

Given the unit processing time per syllable, one can, from the postulates, compute the average time after the beginning of a learning experiment that will pass before any specified syllable is learned. This time, in turn, will give us uniquely the number of errors that will be made with that syllable. While the actual number of errors will be a function of the unit processing time, the percentage that this number represents of the expected total errors is independent of the unit processing time.

The theoretical estimates given in the text tables were obtained as follows: By the postulates, the syllables will be learned in an orderly sequence, each syllable requiring a certain processing time, say \( k \). Each syllable can be identified by its serial order, \( n \), in the list as presented by the experimenter, and also by the order, \( r \), in which it is to be learned. Since learning takes place from both ends of the list, these two orders will not, in general, be identical. Thus, \( n \) is the syllable in order of presentation, may be the same syllable as \( r \), the nth syllable in order of learning. (Technically, the list of syllables in order of learning is a permutation of the original list.) Let \( T_r \) be the time that elapses before the first response to syllable \( n \) that is, until the nth syllable is learned. Then \( T_r = rt \), the number

\[
W_r = m
\]

where \( m \) is a proportionality constant, equal to \( k \) divided by the time per trial.

The numerical value of \( m \) is a function, of course, of the difficulty of the items and the rate of presentation. However, we are only concerned with the fraction of total errors made on a given syllable, and this fraction is clearly independent of \( m \). For \( W_r \) the total number of errors, summed over all syllables, and \( W_r = \frac{W_r}{n} \), the fraction of total errors made on the nth syllable.

\[
w_r = m \sum n = \frac{m}{2} \left( \frac{n^2}{n+1} \right)
\]

where \( n \) is the number of syllables in the list. Suppose, for example, that we are dealing with a list of twelve syllables, then \( \frac{W_r}{m} = 78 \). Hence, the fraction of total errors that will be made on the fourth syllable will be \( \frac{12}{12} \times 0.051 \).

Now, to obtain the serial position curve, we need merely to relabel the syllables from the order in which they are learned to their order of presentation. That is, if \( r \) is the rank in order of learning, of the nth syllable in order of presentation, then the fraction of total errors for the nth syllable will be simply:

\[
w_r = w_r = \frac{r}{n+1}
\]

To apply this result, we must calculate the rank, \( r \), of the nth syllable in order of presentation, as determined by the calculated, \( \frac{r}{n} \). For example, suppose that \( \frac{r}{n} = 0.5 \), and that the immediate memory capacity is 2 syllables and, for simplicity in the calculation, we assume that the items are picked up pairwise in the processing sequence, and stored in the immediate memory for learning. The first two syllables that will be learned (and they must be a rank, \( r = 1 \), and \( r = 1 \), respectively, followed by either the last two syllables on the list, or by syllables three and four, each with probability one-half. The result is that in a list of 12 syllables the third syllable, for example, will have a probability of one-half of being the third syllable in order of learning. A probability of one-quarter of being the fifth syllable, a probability of one-eight of being the seventh, and of one-sixteenth of being ninth or eleven. Averaging these ranks, weighted by their respective probabilities, the average rank of the third

\[
0.755 \text{. The fraction of total errors on the nth syllable will be } w_r = 4.25/38 \frac{r}{n} \text{. (see table 2). All the other predicted values in the tables were computed in the same way.}
\]

The fact that the third syllable has zero probability of being learned fourth, sixth, eighth, etc., is artificially introduced by the calculation, amplification introduced above of handling syllables in pairs. It does not materially affect the serial position curve prediction, as shown by the fact that the computer simulation (which does not use the pairwise learning amplification) generated the same serial position curve prediction.

References


