DOCUMENT RESUME

ED 025 326 PS 001 487

By-Bogartz, Richard S.
Extension of a Theory of Predictive Behavior to Immediate Recall by Preschool Children.
Michigan Univ., Ann Arbor. Center for Human Growth and Development.
Spons Agency-National Inst. of Child Health and Human Development, Bethesda, Md.
Report No-UM-CHGD-8
Pub Date 30 Nov 65
Note-14p.
EDRS Price MF-$0.25 HC-$0.80
Descriptors-Learning Processes, *Mathematical Models, Memory, *Prediction, *Preschool Children, *Probability, Recall (Psychological), *Retention

This paper is concerned with memory functions in sequentially structured behavior. Twenty-five 4- and 5-year-old preschool children participated in a prediction experiment in which a stack of cards (each card alternately having a patch of red or green tape on it) was displayed to the child. The child was presented with a card and asked to predict the color on the next card. Two interval lengths, a long and a short, were used between presentation and prediction. The subject's performance, it was thought, was affected by (1) memory of each trial, (2) effects of the previous response, (3) lagging of attention, (4) guessing, and (5) the variation in interval length. The results from 100 trials indicated that the probability of an error, given a correct response on the previous trial, is greater following the long interval than following the short. It was also found that the probability of an error, given a correct response, is less than or equal to both the probability of a correct response, given an error, and the probability of an error, given an error. The theoretical basis of this task is being used to develop a recall task similar in form to the prediction task. (WD)

Extension of a Theory of Predictive Behavior to Immediate Recall by Preschool Children

Richard S. Bogartz

In the first part of this paper I will describe a theory of young children's behavior in the binary prediction situation and present the results of one test of that theory. I will then describe an experiment which is now in progress in which the prediction situation is transposed into an immediate recall task and indicate how the theory of predictive behavior may be transposed and tested in the recall situation.

The binary prediction or two-choice guessing situation may be paradigmatically represented as a sequence of trials, the nth of which is shown in Fig. 1. The temporal sequence, moving from left to right, consists of a cue of negligible duration (we have used a .3 sec. buzz), a subject-determined latent interval terminated by the predictive response, PR, also taken to be of negligible duration, an experimenter-determined delay interval terminated by onset of the event being predicted, followed by offset at the end of an experimenter-determined exposure interval. Offset of the event terminates the trial and begins an experimenter-determined intertrial interval which is in turn terminated by the cure CUE ? which starts the next trial. An interval which will be of special interest in the discussion to follow is the response-cue interval, the interval between the predictive response on the nth trial and the cue which starts the $n+1$st trial.

When four and five year old preschool children predict in a situation consisting of a sequence of such trials, and when the two events they are predicting, A and B, alternate ABAB...on successive trials, almost all children make errors. These errors do not occur randomly; instead they appear to depend upon the behavior on the previous trial. Putting this more precisely, it appears that for most children the sequence of correct responses and errors approximates a first-order Markov chain with stationary transition probabilities. This means, simply, that the probability of a correct response on a given trial appears to depend only upon whether the subject made a correct response or an error on the previous trial, and that the conditional probabilities characterizing this dependence seem to remain fairly constant over trial sequences of at least 100 trials.

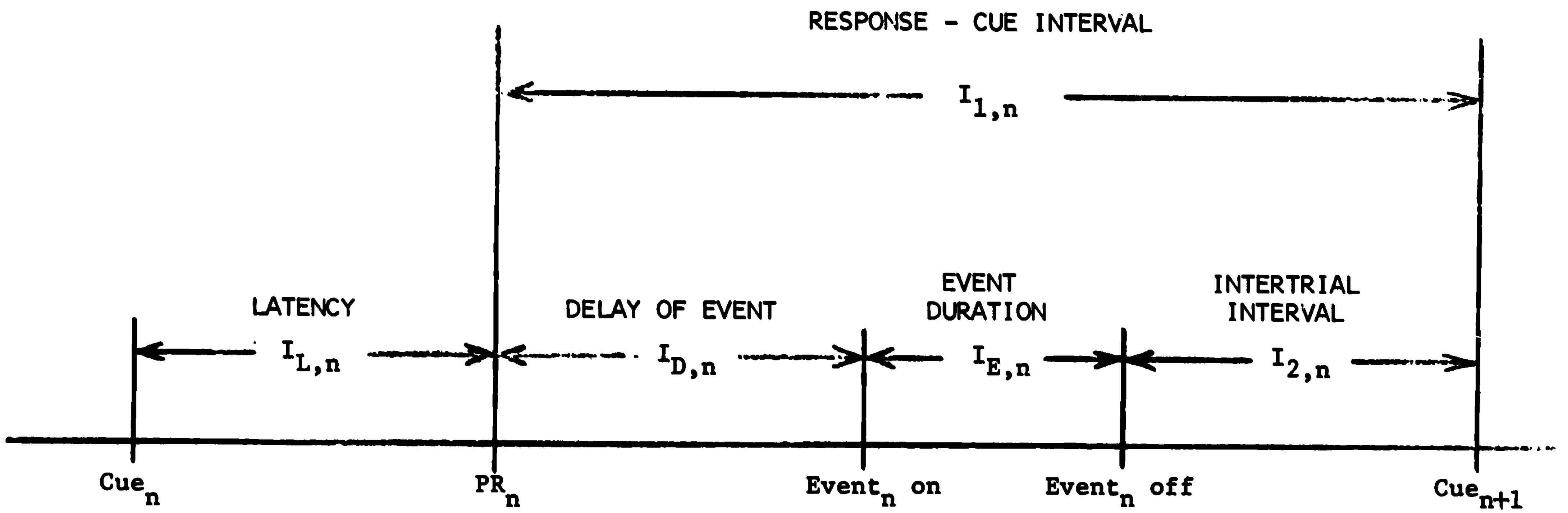

Fig. 1 Paradigmatic representation of the _n_th trial of a two-choice guessing situation.

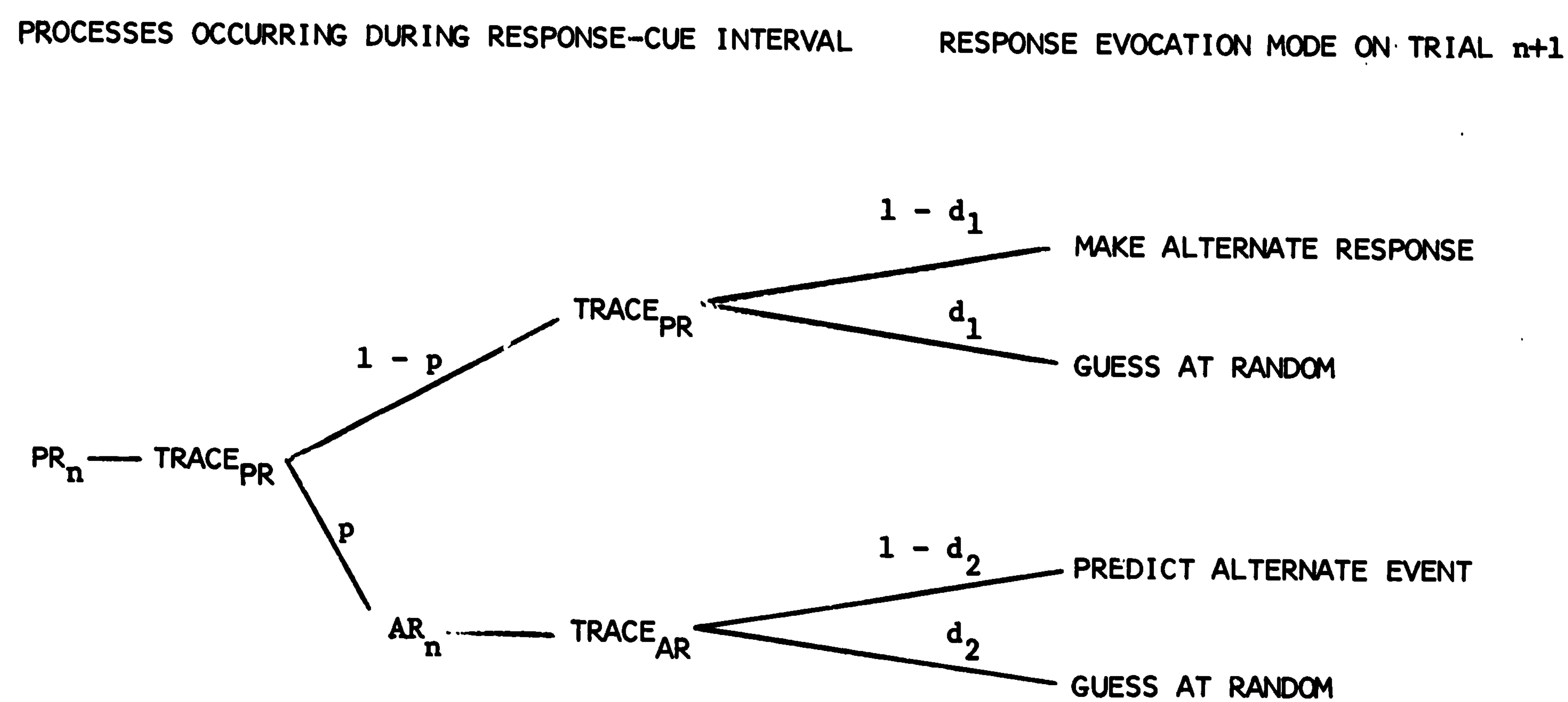

Fig. 2 Summary of the theory of predictive behavior extended to an interpretation of immediate recall.

We have observed that the children do not always seem to be paying attention to the events which they have just predicted. During the interval in which the event is exposed to them they may be looking away and not see it. Such inappropriate orientation sometimes has been observed to coincide with a run of errors in which the sequence of events alternates ABAB...while the child is predicting the events BABA...on the same trials. It seems that the child sometimes just alternates his previous response regardless of the event. We have also observed that the termination of such a run of errors is often preceded by a reorientation to the event during the last error trial, accompanied by a verbal or nonverbal response suggestive of his having noticed that he has made an error.

These observations, in conjunction with Markovian properties of the data, support a theoretical formulation of prediction in the alternation situation which in everyday language can be approximated as follows. On each trial the child guesses the next event at random unless he remembers either the response he made or the event which occurred on the previous trial. If he remembers his previous response, he makes the alternate response; if he remembers the event which occurred, he predicts that the alternate event will occur. In order for him to remember the event which occurred on the previous trial, he must have been paying attention to the event, and if he remembers the event, he does not remember his response.

More precisely, (see Fig. 2) it is assumed that each predictive response produces a stimulus trace which is conditioned to the alternative predictive response. If this trace lasts through the response-cue interval, it elicits the response to which it is conditioned. If, however, the subject makes an attending response to the event, this response furnishes a trace which displaces the trace produced by the predictive response. If the trace produced by the attending response lasts through the intertrial interval, it elicits the response to which it is conditioned, prediction of the alternate event. Either type of trace is assumed to be vulnerable to distracting stimuli which displace the trace, and if no trace is present when the cue occurs, the subject guesses at random.

The consequence of these assumptions is that the sequence of correct responses and errors is a Markov chain with transition matrix where $\alpha = p(1-d_2)$ is the probability of

	C_{n+1}	e_{n+1}
C_n	$1 - \gamma/2$	$\gamma/2$
e_n	$\alpha + \gamma/2$	$\beta + \gamma/2$

elicitation of the response by the event trace, $\beta = (1-p)(1-d_1)$ is the probability of elicitation of the response by the trace of the previous response, and $\gamma = pd_2 + (1-p)d_1$ is the probability of guessing.

Finally, it is assumed that the conditioning states of the traces, that is, the rules which determine which response a given trace will elicit, are determined by preexperimental associations, associations established during pretraining sessions, or learning which is completed during the early trials of the experiment.

In the prediction experiment, 25 four and five year old preschool children served as subjects. Each subject was taken individually to an experimental room and seated opposite the experimenter in front of a small table upon which was a stack of 100 four by six inch white cards concealed behind a small black box which could contain the entire stack. Centered on each card was a 1.5 by 2.0 inch rectangular patch of red or green tape. The colors in the stack were ordered in a simple alternation sequence (RGRG...or GRGR...). The subject was shown the first two cards in the stack, one after the other, after being asked to name the colors on the two cards. Following correct naming of the two colors, a six-volt buzzer was sounded briefly and the child was told that each time he heard the buzzer he was to guess quickly the next color in the stack.

On each of 100 trials, following each buzz, the child made his prediction, the experimenter removed the top card from the stack, turned it color side up in front of the child for about one second, and then placed it color side down in the box. The buzzer sounded for .3 sec. every eight sec., except when the experimenter depressed a foot switch which opened the buzzer circuit. If this occurred, the interbuzz interval was increased from 7.7 secs. to 15.7 secs. A different random sequence of long and short intervals was used with each child; thus approximately half the intervals were long and half were short.

The theory may be applied to the data both with and without inclusion of the interval variable and to group data or to that of an individual subject. When the interval variable is excluded from the analysis, we obtain the maximum likelihood estimates of α,β, and γ which you recall are the probabilities of the three response evocation modes, response alternation, event alternation, and guessing, respectively. Since these three probabilities sum to one, only two need to be estimated from the data. With the two parameter estimates and an estimate of the initial probability of a correct response in hand, it is possible to predict for a group or for an individual subject the values of various sequential statistics which exemplify the extent to which the model summarizes aspects of the data. The sequential statistics which are shown here are: the relative frequency of response triples; the total number of runs of errors; r_j, the number of runs of errors of length j; and C(K), the number of joint occurrences of a correct response on a given trial and a correct response K trials later. There are $2^3 = 8$ response triples, one of which is observed in each group of three adjacent trials. They are: correct, correct, correct; correct, correct, error; correct, error, correct; and so on to error, error, error. A run of errors of length j is simply j consecutive errors immediately preceded and followed by at least one correct response, and the total number of error runs is simply the sum over j of the number of runs of length j. The relative frequency of the triples, the total number of error runs, the number of runs of length j, and the values of C(K) are all known functions of the parameters which I just mentioned and the number of trials in the experiment.

Table 1 shows the group observed and predicted values of the sequential statistics for all 25 subjects and for 21 of the 25. There is good agreement between the observed and predicted values, particularly in view of the fact that although the expected values of the various statistics may be obtained without inclusion of the interval variable, the variances of these statistics are increased by the interval effect. The four excluded subjects exhibited bursts of repetition of the same color prediction which were longer than that which would be compatible with their behavior during the remainder of the trials under the assumptions of the present theory. By repeating the same response on each of a series of trials, they made an error on every other trial, thus obtaining more error runs of length one than expected.

Table 1

Predicted and Observed Values of Various Group Statistics

Statistic	All 25 Subjects Obs.	Pred.	21 of 25 Subjects Obs.	Pred.
3-Tuples				
ccc	.530	.526	.501	.495
cce	.107	.109	.109	.113
cec	.085	.076	.083	.076
cee	.046	.056	.054	.063
ecc	.107	.111	.108	.113
ece	.026	.023	.030	.026
eec	.046	.057	.054	.064
eee	.053	.042	.060	.052
Mean number of runs of errors				
	13.480	13.459	14.048	14.015
Runs of errors, r_j, of length j				
r_1	8.880	7.784	8.619	7.724
r_2	2.000	3.282	2.381	3.467
r_3	1.400	1.384	1.667	1.556
r_4	.680	.583	.810	.698
r_5	.120	.246	.143	.313
Autocorrelation of correct responses, C(K), K trials apart				
C(1)	63.000	62.901	60.286	60.144
C(2)	60.320	58.981	57.238	55.905
C(3)	58.120	57.553	54.714	54.371
C(4)	57.960	56.750	54.714	53.554
C(5)	56.840	56.105	53.476	52.928

Table 2

Predicted and Observed Total Number of Runs of Errors

Subject	Observed	Predicted
1	18	17.24
2	11	11.04
3	17	15.12
4	7	3.66
5	18	18.23
6	17	17.05
7	24	21.85
8	11	10.77
9	13	12.77
10	18	18.10
11	15	14.77
12	18	15.23
13	10	10.17
14	10	10.10
15	18	17.70
16	7	7.03
17	8	8.11
18	11	10.82
19	13	13.02
20	9	4.25
21	9	9.02
22	12	5.93
23	14	14.04
24	12	11.86
25	17	17.08

Table 2 shows the observed number of error runs and predicted number of error runs for the 25 individual subjects. The theory appears to be describing well the behavior of most of the individual subjects as well as that of the group.

Application of the theory to the data with inclusion of the interval variable results in two predictions. The first, a rather weak prediction, is that the probability of guessing following the long interval should be greater than that following the short interval since we expect that the longer the interval, the less likely should it be that either type of trace will still be present and therefore the more likely it should be that the subject will guess. This prediction was confirmed in that the estimate of the probability of a guess following the long interval was .53 and the corresponding estimate for the short interval was .15.

Maximum likelihood estimates of the response evocation mode probabilities for each interval may be used to obtain a sequential statistic which provides a stronger test of the theory with respect to its applicability to the predictive situation when experimenter-controlled sequence of intervals are used. A 3,2-tuple is a joint event consisting of three predictive responses, each of which is either correct or an error, and the two intervening intervals. For example, correct response, short interval, error, long interval, error is a 3,2-tuple; likewise error, long interval, error, short interval, error. In all, there are 32 3,2-tuples which may occur if only two different intervals are used. By estimating two parameters for each interval, it is possible to predict the relative frequencies of all 32 3,2-tuples.

Table 3 shows the observed and predicted values for the 32 3,2-tuples for the data from all 25 subjects. The corresponding table for the 21 subjects showed a slightly better picture.

The data from almost all of the subjects support three implications of the model: the sequence of correct responses and errors is approximately a first order Markov chain with stationary transition probabilities; the probability of an error given a correct response on the previous trial ($\gamma/2$) is less than or equal to both the probability of a correct response given an error ($\alpha + \gamma/2$) and the probability of an error given an error ($\beta + \gamma/2$);

Table 3

Predicted and Observed 3,2-Tuples

3,2-Tuple	Obs.	Pred.	3,2-Tuple	Obs.	Pred.
c S c S c	.156	.164	c L c S c	.137	.130
c S c S e	.016	.013	c L c S e	.006	.010
c S e S c	.009	.008	c L e S c	.032	.028
c S e S e	.004	.006	c L e S e	.018	.023
c S c L c	.123	.130	c L c L c	.114	.103
c S c L e	.034	.047	c L c L e	.031	.037
c S e L c	.011	.008	c L e L c	.033	.030
c S e L e	.004	.006	c L e L e	.021	.021
e S c S c	.025	.030	e L c S c	.030	.032
e S c S e	.003	.002	e L c S e	.004	.003
e S e S c	.011	.014	e L e S c	.013	.013
e S e S e	.016	.012	e L e S e	.014	.011
e S c L c	.027	.024	e L c L c	.025	.025
e S c L e	.011	.009	e L c L e	.009	.009
e S e L c	.011	.015	e L e L c	.011	.014
e S e L e	.014	.011	e L e L e	.008	.010

and the probability of an error given a correct response is greater following the long interval than following the short interval. The data therefore also support the theory which gave rise to these predictions in which the relevant processes are conceptualized in terms of attention, event and response traces, and random guessing.

It should be observed, however, that other conceptions of the subject in terms of different processes could also give rise to the same prediction concerning these properties of the data. Thus, while the goodness of fit is certainly encouraging, its usefulness in establishing the appropriateness of the proposed concepts is limited. The value of identifying model parameters with terms such as "guessing probability" or "the probability of the response-produced trace eliciting the next response" will be established more by the scope of the theory, by its extendability to other behavioral measures and other situations, and by confirmation of predictions, suggested by these identifications, concerning the covariation of parameter values with values of independent variables.

The immediate recall task to which the theory will be extended incorporates two changes directed to delimiting the scope of the theory. Previously, the theory has been applied to prediction of the next event in various binary sequences (including alternation, noncontingent, and Markov sequences). In the recall task the subjects will postdict (recall) the previous event in the ternary sequence of colors Red-Blue-Green-Red-Blue-Green...The events will be colored lights on a panel, with a different color at each vertex of a triangle. The child will be instructed to recall the previous color each time the buzzer sounds. The session will start with a red light, a buzzer will sound, the child will respond, another light will follow, etc. A variable event-cue interval will also be used. The child's response and response latency will be recorded.

With these changes, the sequence of occurrences in the situation will now be: Event-Buzz-Response-Event-Buzz-Response-... as opposed to Buzz-Response-Event-Buzz-Response-Event... in the prediction task. Note, however, that a latecomer to the situation could not tell, even after observing several complete trials, whether he was observing a prediction task or a recall task. In other words, so far as the temporal sequence of observable events is concerned, the two situations are the same once they have been going for a while.

This congruence of the two tasks, in conjunction with the assumption that in the recall task, as in the prediction task, the subject's behavior should be affected by memory of the event, effects of his previous response, lagging of attention, and guessing, suggests at least two models of the situation in terms of the theory described above. The two differ only in the assumption concerning the conditioning state of the trace produced by the recall response. They will be referred to as Models A and B. In both models it is assumed that the absence of any trace results in a guess and that the trace of the event elicits the correct recall response.

In Model A it is assumed that the trace of the previous recall response elicits the name of the next color in the sequence. Thus, if the subject has just said "Red" and he does not pay attention to the next color shown to him, then if the trace of his response is still present when the buzzer sounds, he will say "Blue." This is the sort of assumption to make if we expect that in this situation, as in the alternation prediction situation, the children, when not paying attention, will often give the colors in the right order but out of phase with the events.

In Model B it is assumed that the trace of the recall response elicits that same response. Thus, if the subject has just said "Red," the trace of this response will elicit "Red" on the next trial if it is still present. The idea underlying Model B is that if the children rehearse the name of the color which they have just seen until the buzzer sounds, they may, when attention lags, respond to the trace of their previous response as they would respond to the trace produced by rehearsal. This would result in repetition of their previous response.

To summarize the differences between the two models and to relate Model A to the model used in the prediction situation, it is convenient to discriminate two kinds of errors that can occur in recall. We will say that if the subject's response is one ahead of the correct response in the sequence he has made an e_1 and if it is two ahead he has made an e_2. Thus if Red is correct, "Blue" is an e_1 and "Green" is an e_2. Then the transition matrix implied by Model A is

$$\begin{array}{c|ccc|}
 & c_{n+1} & e_{1,n+1} & e_{2,n+1} \\ \hline
c_n & 1 - 2\gamma/3 & \gamma/3 & \gamma/3 \\
e_{1,n} & \alpha + \gamma/3 & \beta + \gamma/3 & \gamma/3 \\
e_{2,n} & \alpha + \gamma/3 & \gamma/3 & \beta + \gamma/3 \\ \hline
\end{array},$$

and that implied by Model B is

$$\begin{array}{c|ccc|}
 & c_{n+1} & e_{1,n+1} & e_{2,n+1} \\ \hline
c_n & \alpha + \gamma/3 & \gamma/3 & \beta + \gamma/3 \\
e_{1,n} & \alpha + \beta + \gamma/3 & \gamma/3 & \gamma/3 \\
e_{2,n} & \alpha + \gamma/3 & \beta + \gamma/3 & \gamma/3 \\ \hline
\end{array},$$

where in both cases, as in the model for predictive behavior, α, β, and γ are the probabilities that the recall response is elicited by the trace of the event, elicited by the trace of the previous response, and is a guess, respectively. Inspection of the two transition matrices reveals that the difference in assumptions concerning the effect of the response-produced trace is reflected by different equalities and inequalities of various conditional probabilities.

An interesting property of Model A but not of Model B is that the two error states are lumpable. This means that if the type of error is ignored and the responses are simply categorized as correct or error, the sequence of c's and e's will still be a Markov chain with transition matrix

$$\begin{array}{c|cc|}
 & c_{n+1} & e_{n+1} \\ \hline
c_n & 1 - 2\gamma/3 & 2\gamma/3 \\
c_n & \alpha + \gamma/3 & \beta + 2\gamma/3 \\ \hline
\end{array}.$$

This matrix is a special case of the matrix

$$\begin{array}{c|cc|}
 & c_{n+1} & e_{n+1} \\ \hline
c_n & 1 - (N - 1)\gamma/N & (N - 1)\gamma/N \\
c_n & \alpha + \gamma/N & \beta + (N - 1)\gamma/N \\ \hline
\end{array},$$

where N is the number of events in the sequence. Setting N equal to 2 in this matrix we obtain the transition matrix implied by the prediction model. This reflects the fact that Model A is based on the same assumptions concerning the effects of the previous response on the next response as is the prediction model.

In the event that the recall model is compatible with the data, we plan to pursue the problem by (1) using probabilistic ternary sequences in prediction and recall, (2) interpolating distractive or associatively interfering events into the situation at different points in the sequence and evaluating their effects on the parameters as a function of their location in the sequence, (3) following up some very promising results which have supported a prediction concerning a conditional response latency measure which was made using the estimates of α, β, and γ to weight other conditional latencies in a prediction equation generated by the theory plus the auxiliary assumption that the means of three response latency distributions, one associated with each response evocation mode, all exist, and (4) using subjects in relatively small regions of the α-β-γ parameter space in parametric studies of the relationship between certain interval durations and the α, β, and γ values.

In the event that the recall model does not fit, we shall have a lot of recall responses and latencies, and the interesting question of how the instructions to recall teh previous event transform the subject from a Markovian predictor to a non-Markovian postdictor.