APPLICATION OF A STIMULUS SAMPLING MODEL TO CHILDREN'S CONCEPT FORMATION OF BINARY NUMBERS, WITH AND WITHOUT AN OVERT CORRECTION RESPONSE

by

Patrick Suppes and Rose Ginsberg

TECHNICAL REPORT NO. 35
December 14, 1960

PSYCHOLOGY SERIES

INSTITUTE FOR MATHEMATICAL STUDIES IN THE SOCIAL SCIENCES
Applied Mathematics and Statistics Laboratories
STANFORD UNIVERSITY
Stanford, California

APPLICATION OF A STIMULUS SAMPLING MODEL TO CHILDREN'S CONCEPT FORMATION OF BINARY NUMBERS, WITH AND WITHOUT AN OVERT CORRECTION RESPONSE[1]

Patrick Suppes and Rose Ginsberg

In the present experiment, the first of a series of experiments intended to study concept formation in young children, five and six year old Ss are required to learn two fairly difficult concepts each represented by three stimuli. The concepts to be learned are numbers in binary notation and the stimuli representing them are varied by the use of different symbols for the binary digits. A feature of the experimental situation is that it permits the application of a simple sampling model to the data and in addition suggests two methods of analysis, one in terms of paired associate learning and the other, which we shall later refer to as "pure property analysis", more directly concerned with acquisition of the concepts involved.

If, in analyzing the present data the six stimuli representing two concepts are treated as if they were independent, the situation is analagous to a paired associate learning experiment and the data become amenable to analysis in terms of a statistical learning theory that has been successfully applied to paired associate experiments (Bower, 1960).

Having treated all stimuli in the first analysis as if they were independent and having looked at the situation as if it involved learning simply the association between each stimulus and the correct response we

---

1/ The research reported was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

proceed in the second analysis to treat the three stimuli representing the same concept as if they were identical--in effect one stimulus. The learning situation is then reduced to the association of two stimuli and the two appropriate responses and the paired associate model used in the first analysis may again be applied to the rearranged data. In this second method of analysis we view the learning situation as involving the association between the common property of the stimuli which determines a concept and the identification of that concept, which is the correct response. To distinguish the two approaches we designate the first method--six independent stimuli and two responses--the "paired associate" analysis, and the second approach--two independent stimuli and two responses--the "pure property" analysis.

A further aim in the present experiment is to establish the role in young children's learning behavior of an overt correction response following an incorrect response. In two-response situations with adult human Ss when S is informed whether he is correct or incorrect on each trial, the probability of making a correct response is affected by the outcome of a trial rather than the specific response made on that trial. Requiring S to make an overt correction response after reinforcement neither improves his learning rate nor influences his asymptotic behavior (Burke, Estes, Hellyer, 1954). To establish whether an analagous situation obtains when the Ss are young children, half of the 5 year old and half of the 6 year old Ss in the present experiment are required to make an overt correction response after an incorrect response (Correction Group) the remaining Ss are simply told on each trial whether they are correct or incorrect but are not required to make a correction response (Noncorrection Group).

Method

Subjects. There were 48 Ss divided between two groups. Each group consisted of 12 Ss from Grade 1 and 12 Ss from the Kindergarten class of Stanford Elementary School. The average age of the former was 6 years and 3 months, of the latter 5 years and 1 month.

Design. On each trial a stimulus was presented to which S made one of two possible responses. E always indicated whether S's choice on that trial was correct or incorrect. For 24 of the Ss (the Correction Group) a wrong choice was followed by an overt correction response, for the remaining 24 Ss (the Noncorrection Group) no overt correction response was required.

Equipment. The stimuli, printed symbols on 8 1/2" x 11" white cards, were presented one at a time on a small metal reading stand to the front of which a sizable white hook was attached. Two responses were available, each printed in red on a 2 1/2" x 4" card. A one inch composition ring attached to the top of each response card enabled S to hang the chosen response on the reading stand hook.

Stimuli and Responses. Six different stimuli were presented. Three of the stimuli represented 4 in the binary number system, the remaining three represented 5 . The equivalent binary numbers are, of course, 100 and 101 , but as quite small children are often familiar with some combinations of 1 and 0 , different symbols were used in place of the Arabic numerals. The stimuli used and their Arabic numeral equivalents are shown in Table 1.

Insert Table 1 about here.

Table 1

Stimuli and Equivalent Arabic Numerals Used in Learning Trials

Stimuli			Responses
Γ * *	λ π π	↔ Σ Σ	4
Γ * Γ	λ π λ	↔ Σ ↔	5

Sixteen of each stimulus were prepared, making 96 stimuli in all. Four different random sequences were used and matched across groups.

Two responses were available, the Arabic numerals 4 and 5, each response made of 1/4" red tape on a white card. As indicated a composition ring was attached to each response card.

Procedure. The Ss were run individually. Each S was seated on a small chair with the reading stand on a low table in front of him. A response card was placed in front of and to the right and left of the reading stand (the position, left or right, of the responses was balanced across Ss). Ss were told that the numbers on the response cards were 4 or 5. All the children were familiar with these Arabic numerals although recognition was not, of course, necessary to complete the task. It was explained that all the "pictures" to be shown were different ways of writing 4 or 5, that is they were all the "same as" 4 or the "same as" 5. The Ss were instructed to guess whether a "picture" presented on the reading stand was the "same as" 4 or 5 and to indicate their guess by putting the chosen response on the hook in front of the "picture".

A practice card was then placed on the reading stand, two seconds later E said "now" and S placed the chosen response card on the hook. E said "right" if the answer was correct and "no" if the answer was incorrect. In the Noncorrection Group this completed a trial; in the Correction Group, after an incorrect response, the response card was taken from the hook and S given time to place the correct card on the hook. The trial was terminated by E saying "right". In both groups at the end of a trial S was allowed to remove the response card from the hook while E changed stimuli. The correct inter-trial position for the response cards was established before the next stimulus was exposed. A fairly constant speed for each movement on each trial was in general easily established both for S and E. For all Ss it was not necessary after the first few trials for E to initiate S's response by saying "now".

To demonstrate the procedure to S, two practice trials were given. Each of the stimuli used on these trials consisted of two different objects (dissimilar also to the symbols used in the learning trials) and the response to the first stimulus was always judged as correct, to the second stimulus as incorrect. The order of presentation of practice stimuli was reversed for half the Ss across both groups and grades.

The 96 training trials following the practice trials were broken up by a two minute rest period between the first and second 48 trials. The learning trials were succeeded by a second two minute rest period after which 4 test trials were given. The test stimuli, two of which represented 4 and two represented 5, are presented in Table 2 together with the correct responses. No information of any kind was given on these four trials but at the end of this short session all Ss were told they had answered correctly on all trials.

Insert Table 2 about here.

Table 2

Stimuli and Equivalent Arabic Numerals Used in Test Trials

Stimuli		Responses
$\delta\ \epsilon\ \epsilon$	$\alpha\ \Delta\ \Delta$	4
$\delta\ \epsilon\ \delta$	$\alpha\ \Delta\ \alpha$	5

Results and Discussion

Mean Learning Curves. As there was no statistical difference between the percent correct responses for Kindergarten and Grade 1 *Ss* (the total number of correct responses for the two grades being respectively 1526 and 1534) the data from both grades were combined and are presented for Correction and Noncorrection Groups in Figure 1. The *t* of 4.00 computed between overall responses for the two groups was significant at the .001 level.

Insert Figure 1 about here.

Although the Correction Group does significantly better than the Noncorrection Group the *Ss* in the former group do not asymptotically achieve perfect learning. As a matter of interest both groups were divided into *Ss* who reached a criterion of at least 5 correct responses on each of the last three presentations of the 6 stimuli, and *Ss* who did not achieve that criterion. In the Correction Group half of the Grade 1 *Ss* and half of the Kindergarten *Ss* reached criterion, in the Noncorrection Group two Grade 1 and two Kindergarten *Ss* reached criterion. Figure 2 describes the correct responses of Criterion and Noncriterion Groups. The difference in learning rate between *Ss* not reaching criterion in Correction and Noncorrection Groups is still apparently considerably affected by the correction response variable.

Insert Figure 2 about here.

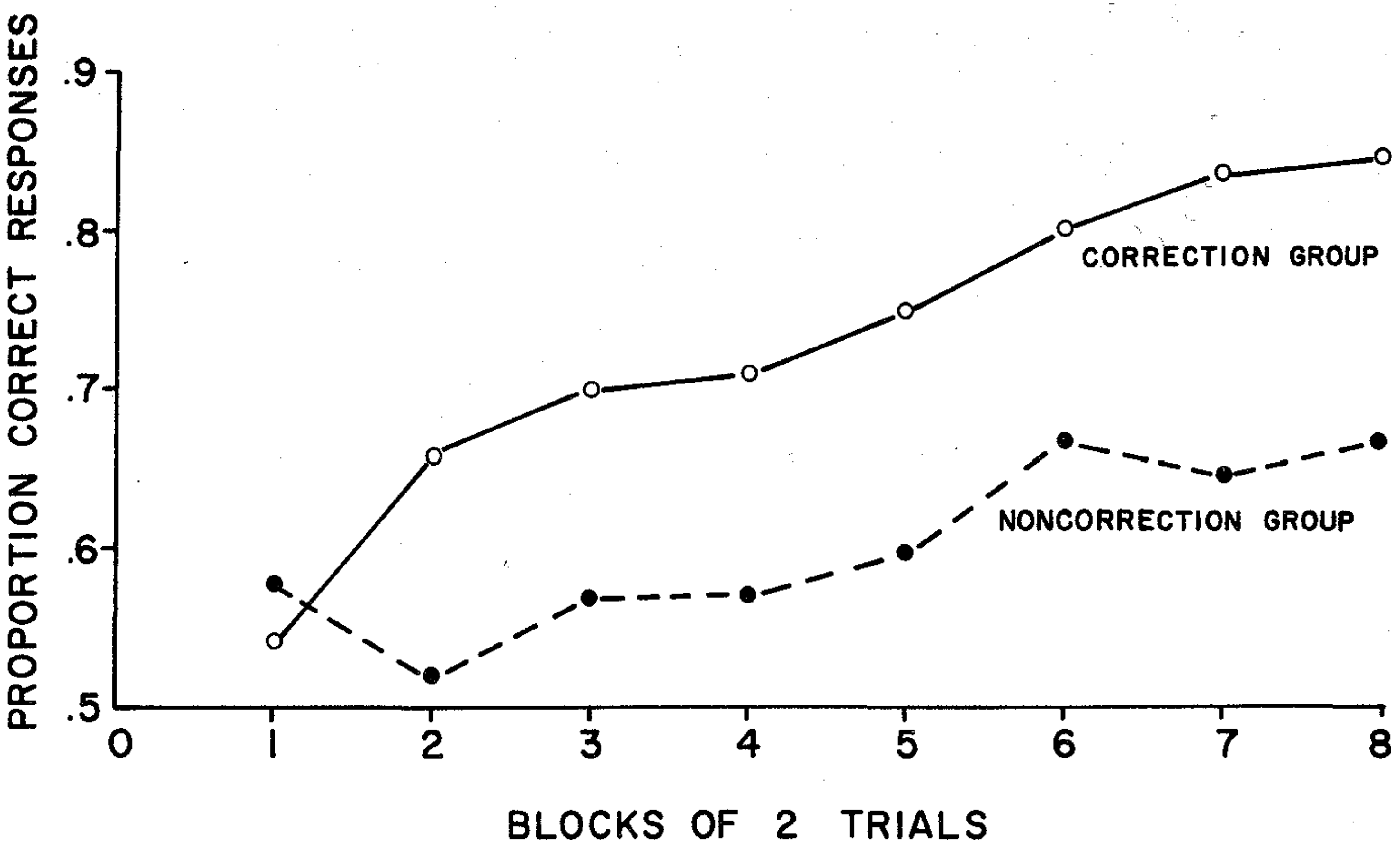

FIGURE 1. Proportion of correct responses.

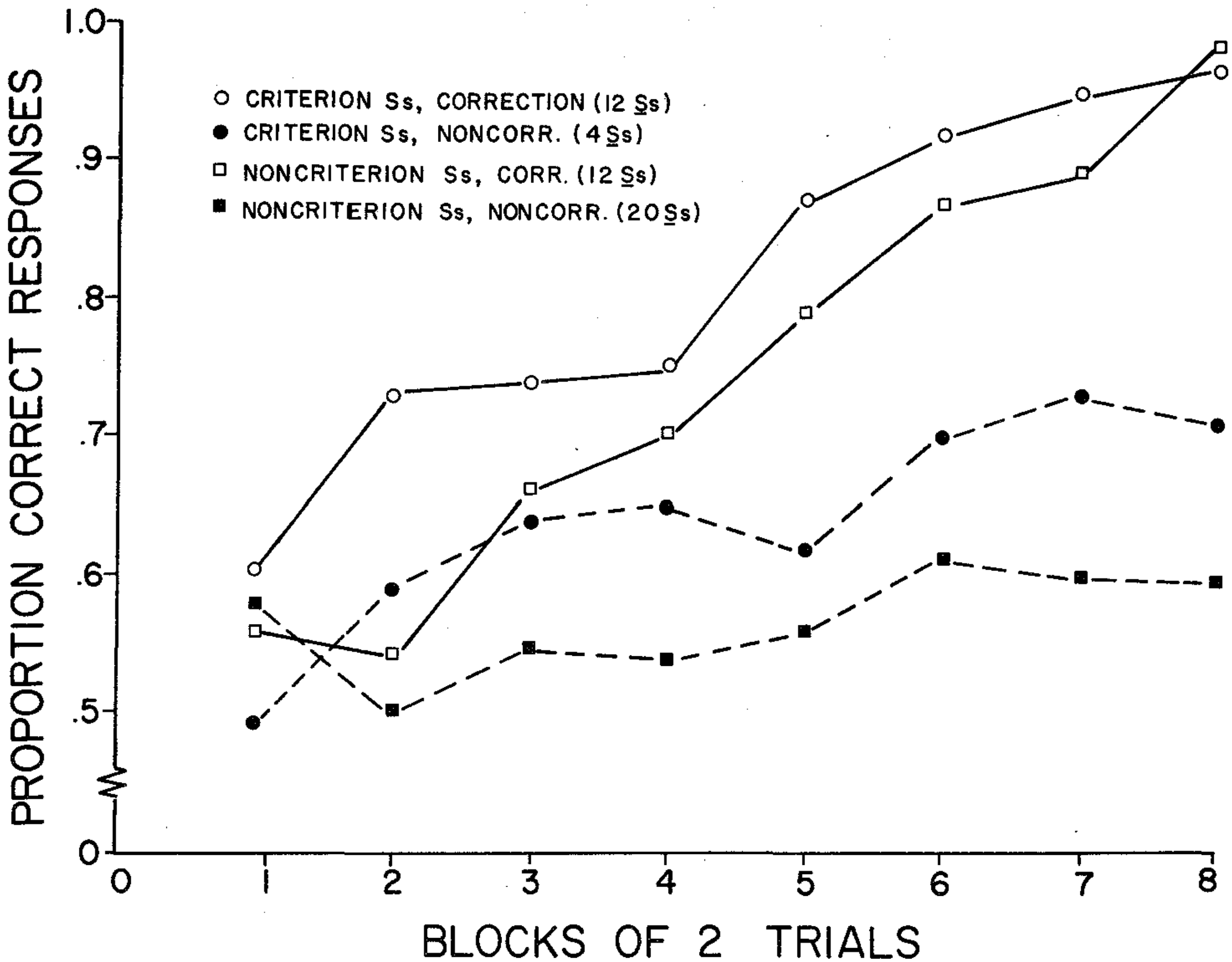

FIGURE 2. Probability of a correct response for Ss with at least 5 out of 6 correct on each of the last three trials (Criterion Ss) and for "Noncriterion" Ss.

A *t* between the overall correct responses of these two subgroups at 2.75 was significant at the .01 level.

Theory. Bowers (1960) has recently presented a one-element model for paired associate learning from which he derived an extensive number of predictions. In effect it is assumed that in a paired associate experiment each item may be represented by one stimulus element which may be in one of two states, conditioned to the correct response or unconditioned. On each trial the stimulus associated with the item presented is sampled. If the stimulus is unconditioned and the trial is reinforced it becomes conditioned to the correct response with probability θ and remains unconditioned with probability $1 - \theta$. At the outset of the experiment the stimulus is assumed to be unconditioned. The probability of a response if an unconditioned stimulus is sampled is $1/k$ where k is the number of available responses. The probability of a correct response if a conditioned stimulus is sampled, is unity.

The statistics which Bower derives are presented in some detail and are not therefore developed here. However, since many of the derived expressions assume a large number of trials or a large value of θ they are not appropriate in the present case where both n and θ are small. Consequently wherever necessary we have re-estimated the statistics for finite n and small θ (see Appendix). At the same time the axioms of this theory require that a reinforcement occur on each trial. In a two-response situation with adult *Ss* the information that a response is incorrect is an effective reinforcement and requiring *S* to make an overt correction response does not improve his rate of learning. From the present results it seems clear that we do not have an analagous situation when the *Ss* are children in that the rate of learning is affected by the addition of an overt correction response. There is, therefore, some doubt that the reinforcement axioms of Bower's model are appropriate to the Noncorrection situation and the model should, therefore, strictly be applied only to the Correction Group of the present experiment. As a matter of interest the statistics of the paired associate analysis have been evaluated for both groups.

Paired Associate Analysis. Using the one-element model described, with each stimulus treated as an independent item, the single parameter θ was estimated from the number of errors. The value of θ for the Correction Group was 0.088, for the Noncorrection Group, 0.031. The theoretical and empirical mean error curves plotted over blocks of 2 trials are presented in Figure 3.

Insert Figure 3 about here.

A more efficient estimation of θ which makes use of more of the data can be made rather easily with the present model, from the joint probabilities. Where we designate the probability of an i^{th} response on trial *m* and a j^{th} response on trial *m+1*, i.e., $P\{A_{i,m}, A_{j,m+1}\}$ as $P_m(ij)$ and $i,j = 1,2$ (where "1" denotes the correct response, "2" the incorrect response), it can be shown that:

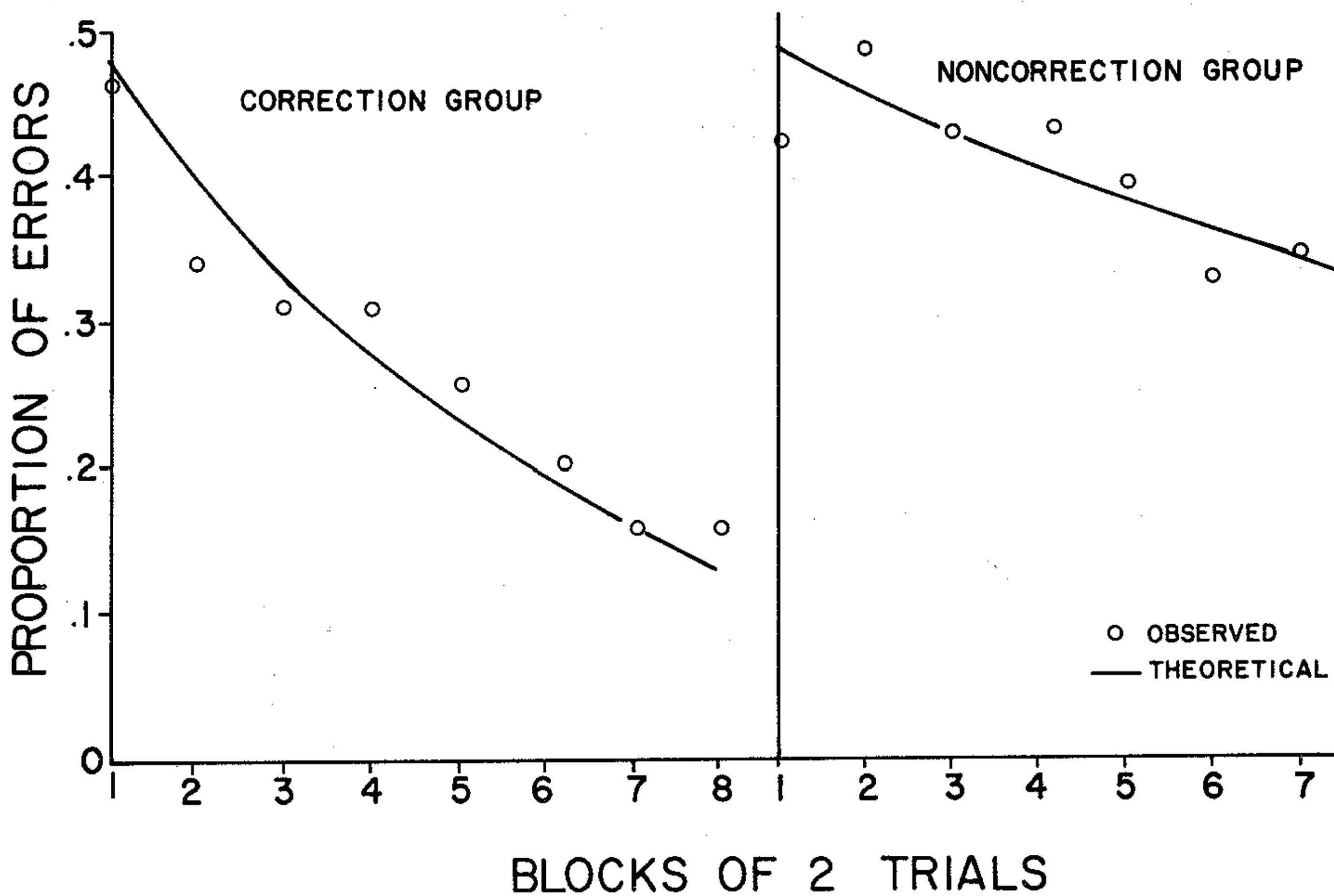

FIGURE 3. Theoretical and observed errors. Paired associate analysis.

$$(1) \quad P_m(11) = 1 - (1-\theta)^{m-1}(3/4 - \theta/4)$$

and

$$(2) \quad \overline{P}(11) = 1/(n-1) \sum_{m=1}^{n-1} [1 - (1-\theta)^{m-1}(3/4 - \theta/4)]$$

$$= 1 - (3/4 - \theta/4)\left[\frac{1-(1-\theta)^{n-1}}{(n-1)\theta}\right]$$

and similarly

$$(3) \quad \overline{P}(12) = \overline{P}(22) = (1-\theta)\left[\frac{1-(1-\theta)^{n-1}}{4(n-1)\theta}\right] ,$$

$$(4) \quad \overline{P}(21) = (1+\theta)\left[\frac{1-(1-\theta)^{n-1}}{4(n-1)\theta}\right] .$$

The above $\overline{P}(ij)$ values, calculated with θ obtained from the mean error curve, are presented in Table 3 together with the observed $\overline{P}(ij)$. The θ values estimated from each of the equations (1) - (4) are also given in Table 3. A least squares fit over the four θ's estimated from the $\overline{P}(ij)$'s

---

Insert Table 3 about here.

---

of the Correction Group yields a value of 0.089 which is surprisingly close to the mean error curve estimate of 0.088 .

Table 4 presents some further predicted and observed statistics for both Correction and Noncorrection Groups. For the former group the model, judged from the statistics presented, describes the data rather well. The least accurate predictions are the sequential statistics--autocorrelation of

Table 3

Average Joint Response Probabilities, $\bar{P}(ij)$

	Correction Group			Noncorrection Group	
$\bar{P}(ij)$	Predicted	Observed	θ Estimated From $\bar{P}(ij)$	Predicted	Observed
$\bar{P}(11)$	0.587*	0.589	0.089	0.399**	0.394
$\bar{P}(12)$	0.129	0.131	0.086	0.196	0.202
$\bar{P}(21)$	0.155	0.156	0.086	0.209	0.212
$\bar{P}(22)$	0.129	0.124	0.094	0.196	0.192

* Estimated with $\theta = 0.088$

** Estimated with $\theta = 0.031$

errors and the difference between the joint probabilities of error following error and success following error--and it is precisely in these statistics that the lack of independence between stimuli would be expected to show up. For the Noncorrection Group as we have indicated there is considerable doubt that the basic reinforcement axioms of the one-element model make it appropriate to the situation of this group and we do not necessarily expect this model to describe the Noncorrection Group data. From Table 4 it can be seen that the fit nowhere approaches that of the Correction Group. On the other hand, none of the observed values depart strikingly from those predicted.

Insert Table 4 about here.

On the whole we must conclude that the model outlined describes much of the data of the Correction Group very well indeed.

Pure Property Analysis. The one-element model described was again applied to the data, but in this instance it was assumed that the stimuli representing a single concept could be treated as a single stimulus. For example, there were three items to which the correct response was the Arabic numeral *4*. These three items were treated as if they were identical, representing one stimulus to which *S* must learn to associate the response *4*. Therefore, instead of six items as in the previous analysis we now had only two items--each to be associated with one of the two responses available. The one-element model as described above was used to analyze the data in this form for the Correction Group only.

Table 4

Predicted and Observed Statistics for Correction and Noncorrection Groups

Paired Associate Analysis

	Statistic	Correction Group		Noncorrection Group	
		Predicted	Observed	Predicted	Observed
1.	Expected errors per item per S		4.38		6.37
2.	s.d. of errors per item per S	3.13	3.00	3.05	2.97
3.	Expected errors before 1st success	0.92	0.94	0.97	0.89
4.	s.d. of errors before first success	1.30	1.29	1.36	1.44
5.	Expected number of success runs	2.82	2.83	3.65	3.67
6.	Expected number of error runs	2.44	2.48	3.42	3.53
7.	Expected error runs of length 1	1.36	1.48	1.85	2.00
8.	Expected error runs of length 2	0.59	0.52	0.86	0.78
9.	Expected error runs of length 3	0.27	0.20	0.40	0.41
10.	Expected number of alternations of success and failure	4.25	4.31	6.07	6.20
11.	Expected errors on trial $\underline{n}$ and on trial $\underline{n+k}$				
	(a) k = 1	1.94	1.86	2.94	2.87
	(b) k = 2	1.71	1.51	2.69	2.67
12.	Net difference between success following error and error following error	0.375	0.486	0.188	0.306
13.	Net difference between error following error and error following success	0.000	0.111	0.000	0.160
14.	$P(A_{1,n} \mid A_{2,n-1})$	0.544	0.558	0.516	0.525

The single parameter θ estimated from the number of errors, was 0.0284 for the Correction Group. As the present analysis is concerned with two distinct items as compared with the six items of the paired associate analysis it follows, of course, that the method of estimating θ from the mean errors will give a parameter value one-third that of the estimate for the paired associate approach. The theoretical and empirical curves for the Correction Group, plotted over blocks of four trials, are presented in Figure 4.

Insert Figure 4 about here.

Table 5 lists some of the obtained and predicted statistics for the data analyzed for pure property learning. The approximate fit of the model is not bad, but it does not compare with the fit of the paired associate model.

Insert Table 5 about here.

In view of the excellent fit of the paired associate model to the present data and the comparative lack of success of the pure property model the immediate question is whether the experimental situation, in fact, involved concept formation learning or was reduced by the difficulty of the task to a paired associate situation. If the <u>S</u>s were learning a concept rather

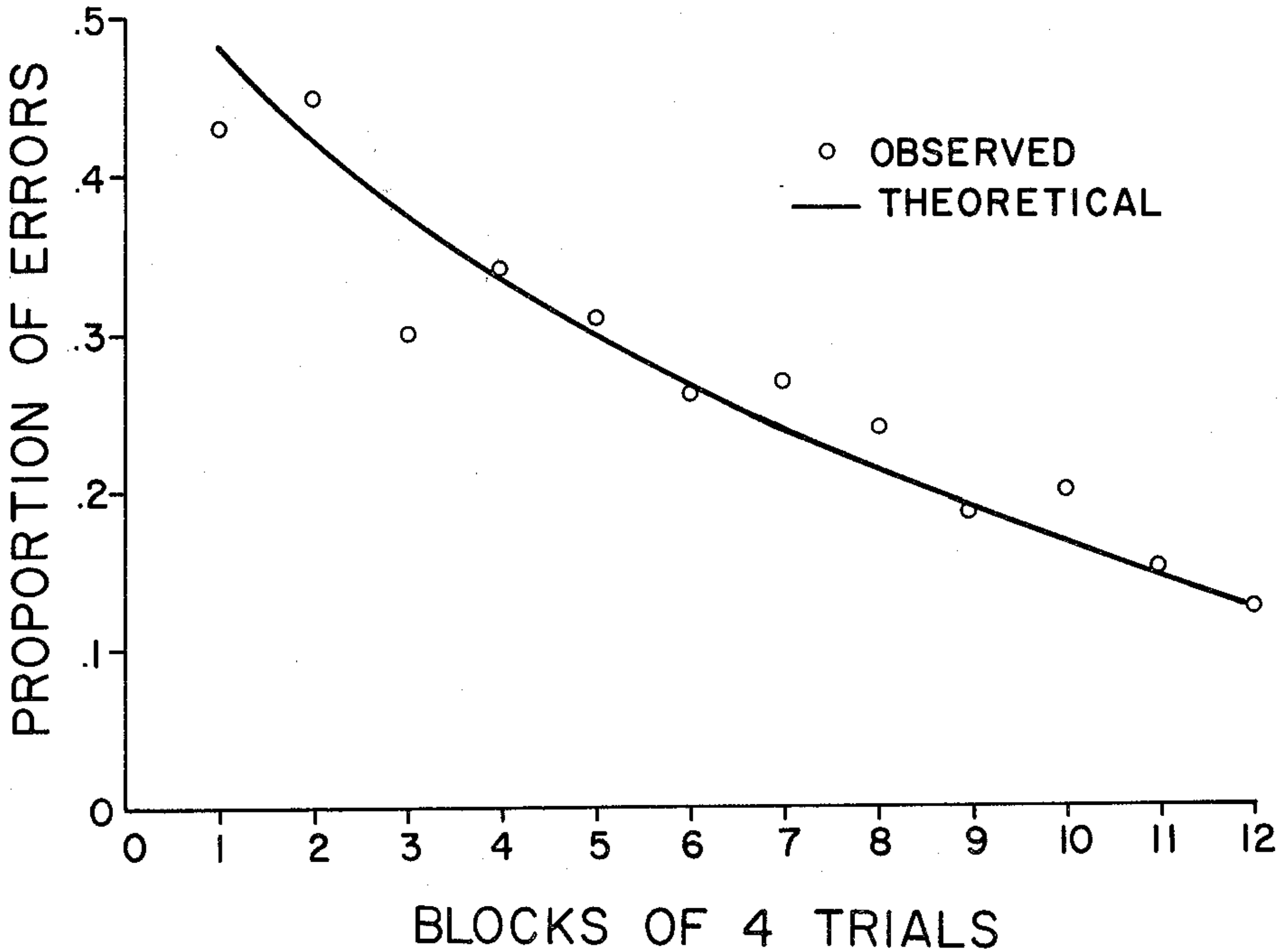

FIGURE 4. Theoretical and observed errors. Pure property analysis for correction group.

Table 5

Predicted and Observed Statistics. Correction Group.

Pure Property Analysis

	Statistic	Predicted	Observed
1.	Expected errors per item per S		13.12
2.	s.d. errors per item per S	8.85	6.29
3.	Expected errors before first success	0.972	1.06
4.	s.d. errors before the first success	1.36	1.49
5.	Expected number of success runs	7.28	8.75
6.	Expected number of error runs	6.78	8.50
7.	Error runs of length 1	5.76	5.57
8.	Expected number of alternations successes and failures	13.06	16.25
9.	Expected errors on trial n and on trial n+k, k = 1	6.34	4.48

than, or in addition to, association of specific stimulus and response, the probability of a correct response when S is faced with a completely new stimulus illustrating the same concept should be at least greater than chance. For both Correction and Noncorrection Groups the initial probability of a correct response on the first experimental trial of the original learning session was 0.5 . Over the four test trials introduced after the 96 learning trials, $P(A_1)$ for the Correction Group was 0.645 , for the Noncorrection Group 0.479 . The obtained value for the Correction Group is significantly different from chance at the .01 level (t of 2.76 with 23 df). Clearly, insofar as the Correction Group was concerned, although the concept was not perfectly learned, acquisition was at least well on the way.

We cannot, at this point, attempt to reconcile the success on the one hand of a paired associate model in describing the present experiment and the evidence, on the other hand, that the Ss were at least in process of forming a concept. The difficulty of the experimental task evidenced by the less than perfect performance on the final acquisition trials probably prejudiced our results in favor of a paired associate approach, but would fail to account for the increased probability of a correct response over the test trials. The adequacy of the paired associate model in a situation in which the majority of Ss acquire the concepts presented remains to be tested.

We do remark that none of the above data directly test the independence assumption of the paired associate model. With a small number of items per concept it is possible for the average data given in Table 4 to satisfy this model and yet for the response to different items to be statistically dependent. An experiment with a sequential presentation of items designed directly to test the independence assumption is now underway.

Summary

Five and six year old Ss were required to learn two concepts in the binary number system, each concept represented by three different stimuli. The Ss were divided into two groups--in the Correction Group each S was required to make an overt correction response after an incorrect response, in the Noncorrection Group the S was simply informed at the end of a trial whether he had responded correctly or incorrectly. The results indicate that the group making an overt correction response performed significantly better than the Noncorrection Group.

Two methods of analysis were used. For the paired associate analysis the stimuli were treated as if they were independent items and for the pure property analysis, all stimuli describing a single concept were treated as if they were identical items. A simple sampling model for paired associate learning was applied to the data in each case. For the paired associate analysis the fit of the model to the data of the Correction Group was very good but was not as good for the Noncorrection Group where it is doubtful that the situation accords with the basic reinforcement axioms of the model. The pure property analysis was performed only with the data of the Correction Group and did not approach the adequacy of the paired associate analysis.

Test trials on new stimuli following the learning trials provided evidence that the _Ss_ were learning the concept involved. It was pointed out that evidence of concept learning is not incompatible with success of a paired associate analysis in the present experimental situation which does not, in fact, provide a test of the independence assumption of the model.

References

Bower, G.H. Properties of the one-element model as applied to paired associate learning. Technical Report No. 31, Contract Nonr 225(17), Institute for Mathematical Studies in the Social Sciences, Applied Mathematics and Statistics Laboratories, Stanford University, 1960.

Burke, C.J., Estes, W.K., and Hellyer, S. Rate of verbal conditioning in relation to stimulus variability. J. exp. Psychol., 1954, 48, 153-161.

Appendix

The derivations presented below are in general those in which consideration of a finite number of trials entails considerable modification of the theoretical predictions presented in Technical Report No. 31 (1960). For each statistic the appropriate referent number in Report 31 will be indicated by a number on the right enclosed in square brackets, and any equations used in the following derivations which have been taken directly from the Report, will also be identified in the same manner.

The arguments for each statistic will not be restated here and unless otherwise indicated the steps of each derivation will be those originally laid out.

Definitions. 1. X_m is a response random variable which takes on the value 1 if an error occurs on trial $\underline{m}$ and 0 if a success occurs on trial $\underline{m}$. $m = 1, \ldots, n$

2. N is the number of available experimental responses. In the present experiment $N = 2$.

<u>Total Errors</u>

The expected total number of errors, per item, per subject, u_1, is given by

(1) $$u_1 = \sum_{1}^{n} (1-\frac{1}{N})(1-\theta)^{m-1} = \frac{(1-\frac{1}{N})[1-(1-\theta)^n]}{\theta}$$ [7a]

The variance of the total errors,

(2) $$\mathrm{Var}\left[\sum_{m=1}^{n} X_m\right] = \sum_{m=1}^{n} \mathrm{Var}(X_m) + 2 \sum_{1 \le i < j} \mathrm{cov}(X_i X_j) \quad . \qquad [13]$$

To get the first term on the right of (2)

(3) $$\sum_{m=1}^{n} \mathrm{Var}(X_m) = \sum_{m=1}^{n} (1-\tfrac{1}{N})(1-\theta)^{m-1}[1 - (1-\tfrac{1}{N})(1-\theta)^{m-1}] \qquad [14]$$

$$= u_1 - \sum_{m=1}^{n} (1-\tfrac{1}{N})^2 (1-\theta)^{2(m-1)}$$

$$= u_1 - (1-\tfrac{1}{N})^2 \frac{[1-(1-\theta)^{2n}]}{1-(1-\theta)^2}$$

$$= u_1[1 - (1-\tfrac{1}{N}) \frac{1+(1-\theta)^n}{2-\theta}] \quad . \qquad [15]$$

To get the covariance term we note,

(4) $$\sum_{i<j} \mathrm{Cov}(X_i X_j) = \sum_{j=2}^{n} \sum_{i=1}^{j-1} (1-\tfrac{1}{N})^2 (1-\theta)^{j-1}[1 - (1-\theta)^{i-1}] \qquad [19]$$

$$= (1-\tfrac{1}{N})^2 \sum_{j=2}^{n} (j-1)(1-\theta)^{j-1} - (1-\tfrac{1}{N})^2 \sum_{j=2}^{n} (1-\theta)^{j-1} \sum_{i=1}^{j-1} (1-\theta)^{i-1} \quad .$$

The first term on the right of (4) is

$$(5)\quad (1-\frac{1}{N})^2 \sum_{j=2}^{n} (j-1)(1-\theta)^{j-1} = (1-\frac{1}{N})^2 [\sum_{j=1}^{n} j(1-\theta)^{j-1} - \sum_{j=1}^{n} (1-\theta)^{j-1}]$$

$$= (1-\frac{1}{N})^2 \sum_{j=0}^{n} j(1-\theta)^{j-1} - \frac{1-(1-\theta)^n}{\theta}$$

$$= (1-\frac{1}{N})^2 \left[\frac{d}{d(1-\theta)} \frac{1-(1-\theta)^{n+1}}{1-(1-\theta)} - \frac{1-(1-\theta)^n}{\theta}\right]$$

$$= (1-\frac{1}{N})^2 \left[\frac{-[1-(1-\theta)](n+1)(1-\theta)^n+1-(1-\theta)^{n+1}}{[1-(1-\theta)]^2} - \frac{1-(1-\theta)^n}{\theta}\right]$$

$$= (1-\frac{1}{N})^2 \left[\frac{n(1-\theta)^{n+1}-(n+1)(1-\theta)^n+1}{\theta^2} - \frac{1-(1-\theta)^n}{\theta}\right]$$

and the last term of (4) is

$$(6)\quad (1-\frac{1}{N})^2 \sum_{j=2}^{n} (1-\theta)^{j-1} \sum_{i=1}^{j-1} (1-\theta)^{i-1} = (1-\frac{1}{N})^2 \sum_{j=2}^{n} \left[(1-\theta)^{j-1} \frac{[1-(1-\theta)^{j-1}]}{\theta}\right]$$

but when $j = 1$, $[1-(1-\theta)^{j-1}] = 0$, so that we can sum from $j = 1$ and

$$(7)\quad (1-\frac{1}{N})^2 \sum_{j=1}^{n} \left[(1-\theta)^{j-1} \frac{[1-(1-\theta)^{j-1}]}{\theta}\right] = \frac{(1-\frac{1}{N})^2}{\theta} \left[\frac{1-(1-\theta)^n}{\theta} - \frac{1-(1-\theta)^{2n}}{\theta(2-\theta)}\right].$$

Putting together (5) and (7) we find that,

$$(8)\quad \sum_{i<j} \mathrm{Cov}(X_iX_j) = \frac{(1-\frac{1}{N})^2}{\theta}\left[\frac{n(1-\theta)^{n+1}-n(1-\theta)^n-(1-\theta)^n+1}{\theta} - [1-(1-\theta)^n] - \frac{1-(1-\theta)^n}{\theta} + \frac{1-(1-\theta)^{2n}}{\theta(2-\theta)}\right]$$

$$= \frac{(1-\frac{1}{N})^2}{\theta}\left[\frac{1-(1-\theta)^{2n}}{\theta(2-\theta)} - 1 - (n-1)(1-\theta)^n\right] .$$

From the results of (3) and (7), we obtain

$$(9)\quad \mathrm{Var}(\Sigma X_m) = u_1\left[1-(1-\frac{1}{N})\frac{1+(1-\theta)^n}{2-\theta}\right] + \frac{2(1-\frac{1}{N})^2}{\theta}\left[\frac{1-(1-\theta)^{2n}}{\theta(2-\theta)} - 1 - (n-1)(1-\theta)^n\right] . \qquad [22]$$

Errors before the first success, $E(F)$ are,

$$(10)\quad E(F) = (1-\frac{1}{N})(1-\alpha)\sum_{k=1}^{n} k\alpha^{k-1} \qquad [41]$$

$$= (1-\frac{1}{N})(1-\alpha)\,\frac{d}{d(\alpha)}\,\frac{1-\alpha^{n+1}}{1-\alpha}$$

$$= (1-\frac{1}{N})\left[\frac{n\alpha^{n+1}-(n+1)\alpha^n+1}{(1-\alpha)}\right] \simeq \frac{1-\frac{1}{N}}{1-\alpha} ,$$

where $\alpha = (1-\frac{1}{N})(1-\theta)$, and similarly

$$(11)\quad \mathrm{Var}(F) \cong E(F)[1+(1-2\theta)E(F)] . \qquad [41]$$

<u>Theorems about runs of errors</u>

The number of j-tuples of errors, $E(u_j)$, are

(12) $$E(u_j) = \alpha^{j-1} \sum_{m=1}^{n-(j-1)} (1-\frac{1}{N})(1-\theta)^{m-1}$$ [29]

$$= \alpha^{j-1}(1-\frac{1}{N})\left[\frac{1-(1-\theta)^{n-(j-1)}}{\theta}\right].$$

The total number of error runs, $E(R)$,

(13) $$E(R) = u_1 - (1-\frac{1}{N})\alpha\left[\frac{1-(1-\theta)^{n-1}}{\theta}\right] .$$ [31]

Error runs of length "j" , $E(r_j)$,

(14) $$E(r_j) = E(u_j) - 2E(u_{j+1}) + E(u_{j+2})$$ [31]

$$= (1-\frac{1}{N})\alpha^{j-1} \frac{1-(1-\theta)^{n-(j-1)}}{\theta} - \frac{1-(1-\theta)^{n-j}}{\theta} \alpha^j + (1-\frac{1}{N})\alpha^{j+1} \frac{1-(1-\theta)^{n-j-1}}{\theta} .$$

The total number of success runs, $E(S)$,

(15) $E(S) = E(V_1) - E(V_2)$ where V_j = j-tuples of successes [30]

(16) $$E(V_1) = 1 - u_1$$

(17) $$E(V_2) = \sum_{m=1}^{n-1} \Pr\{X_m = 0\}\Pr\{X_{m+1} = 0 | X_m = 0\}$$ [29]

where $\Pr\{X_m = 0\} = 1 - (1-\frac{1}{N})(1-\theta)^{m-1}$ [6]

and $\Pr\{X_{m+1} = 0 | X_m = 0\} = 1 - \dfrac{\frac{1}{N}(1-\theta)^{m-1}[(1-\theta)(1-\frac{1}{N})]}{1-(1-\frac{1}{N})(1-\theta)^{m-1}}$ [35]

so

$$E(V_2) = \sum_{m=1}^{n-1} \{1 - (1-\frac{1}{N})(1-\theta)^{m-1} - \frac{1}{N}(1-\theta)^{m-1}[(1-\theta)(1-\frac{1}{N})]\}$$

$$= (n-1) - [1+\frac{1}{N}(1-\theta)](1-\frac{1}{N})\left[\frac{1-(1-\theta)^{n-1}}{\theta}\right] .$$

Putting together the results of (16) and (17), we obtain

(18) $$E(S) = 1 - u_1 - (n-1) + [1+\frac{1}{N}(1-\theta)](1-\frac{1}{N})\left[\frac{1-(1-\theta)^{n-1}}{\theta}\right] .$$

The number of alternations of successes and failures, (A_m) ,

(19) $$E(A_m) = Pr\{X_{m+1} = 1 | X_m = 0\}Pr\{X_m = 0\} + Pr\{X_{m+1} = 0 | X_m = 1\}Pr\{X_m = 1\} \quad [37]$$

$$= \frac{1}{N}(1-\theta)^{m-1}[(1-\theta)(1-\frac{1}{N})] + [1 - (1-\theta)(1-\frac{1}{N})](1-\frac{1}{N})(1-\theta)^{m-1}$$

$$= (\frac{2\alpha-1}{N} + 1 - \alpha)(1-\theta)^{m-1} .$$

(20) $$E[\sum_{m=1}^{n-1} A_m] = (\frac{2\alpha-1}{N} + 1 - \alpha)\frac{1-(1-\theta)^{n-1}}{\theta} . \quad [39]$$

Expected errors on trial <u>m</u> and on trial <u>m + k</u> , $E(C_{k,m})$,

(21) $$E(C_{k,m}) = (1-\frac{1}{N})(1-\theta)^k(1-\frac{1}{N})(1-\theta)^{m-1} . \quad [10]$$

(22) $$C_k = \sum_{m=1}^{n-k} C_{k,m} = (1-\frac{1}{N})^2(1-\theta)^k[\frac{1-(1-\theta)^{n-k}}{\theta}] . \quad [11]$$

Difference between success following error and error following error, (H_m),

$$H_m = (1-X_{m+1})(X_m) - X_{m+1}X_m \quad .$$

(23) $E(H_m) = Pr\{X_{m+1}=0 \mid X_m=1\}Pr\{X_m=1\} - Pr\{X_{m+1}=1 \mid X_m=1\}Pr\{X_m=1\}$

where $Pr\{X_{m+1}=1 \mid X_m=1\} = \alpha$ [28]

so

(24) $$E(H_m) = (1-\alpha)(1-\frac{1}{N})(1-\theta)^{m-1} - \alpha(1-\frac{1}{N})(1-\theta)^{m-1}$$

(25) $$E(H) = E\left[\sum_{m=1}^{n-1} H_m\right] = (1-\frac{1}{N})(1-2\alpha)\left[\frac{1-(1-\theta)^{n-1}}{\theta}\right] .$$

Joint Probabilities

Let C_m be the probability that an element is conditioned at the beginning of trial m. Then

$$C_m = 1-(1-\theta)^{m-1}$$

and

(26) $$P_m(11) = Pr\{A_{1,m}A_{1,m+1}\} = C_m + (1-C_m)(\frac{\theta}{2}+\frac{1-\theta}{4})$$

$$= 1-(1-\theta)^{m-1}(\frac{3}{4}-\frac{\theta}{4})$$

(27) $$\overline{P}(11) = \frac{1}{n-1} \sum_{1}^{n-1} \left[1 - (1-\theta)^{m-1} \left(\frac{3}{4} - \frac{\theta}{4}\right) \right]$$

$$= 1 - \left(\frac{3}{4} - \frac{\theta}{4}\right) \frac{1-(1-\theta)^{n-1}}{(n-1)\theta}$$

(28) $$\overline{P}(12) = \overline{P}(22) = \frac{1}{n-1} \sum_{1}^{n-1} (1-\theta)^{m-1} \frac{1-\theta}{4}$$

$$= \frac{(1-\theta)[1-(1-\theta)^{n-1}]}{4(n-1)\theta}$$

(29) $$\overline{P}(21) = \frac{1}{n-1} \sum_{1}^{n-1} (1-\theta)^{m-1} \left[\frac{\theta}{2} + \frac{1-\theta}{4}\right]$$

$$= \frac{(1+\theta)[1-(1-\theta)^{n-1}]}{4(n-1)\theta}$$