JOURNAL OF MATHEMATICAL PSYCHOLOGY: 1, 143-169 (1964)

# Component Learning in the Four-Category Concept Problem[1]

TOM TRABASSO

*University of California, Los Angeles, California*

AND

GORDEN BOWER

*Stanford University, Stanford, California*

The four-category concept identification task is analyzed as involving two single-cue subproblems. It is supposed that the two concurrent subproblems are worked on independently by the subject, and further, that subproblem learning is a probabilistic, all-or-nothing event. Various lines of evidence are drawn from the data to document these assertions. In particular, the theory successfully predicts the behavior of subjects in the four-category task by using parameters estimated from other subjects learning only single-cue subproblems.

One rationale for the development of theories based upon simple behavioral situations is the promise that the model so obtained can be extended to elucidate behavior in more complex situations. This paper exploits this approach and attempts to extend a theory of elementary concept learning (Bower and Trabasso, 1963a) to handle a slightly more complex case.

In concept identification (CI), the subject is shown a series of complex stimulus patterns, usually geometric figures, which vary in $n$ binary dimensions (color, size, shape, etc.). The $2^n$ patterns are divided into several mutually exclusive and exhaustive classes. To each pattern, the subject attempts to anticipate the correct classification; following his response, he is informed of the correct answer.

In two-category CI a single dimension is relevant. For example, if color (orange or

[1] This work was performed at Stanford University while the first author was supported by a Public Health Service postdoctoral fellowship, 5 F2 MH-18, 070-02, from the National Institute of Mental Health, Public Health Service. Financial support for this research was also provided by grants, M-3849 and MH 07610-01, to the second author. Acknowledgement is made to Milton Kielsmeier for his assistance in providing research facilities at Foothill Junior College.

blue) is relevant, orange patterns are assigned to one category, and blue patterns to the other. The theory developed earlier for the two-category problem (Bower and Trabasso, 1963a) assumes that the subject is sampling dimensions to test for their relevance to the solution. The problem is solved "suddenly" on that trial when the subject begins to attend consistently to the relevant dimension which determines the classification. The evidence for insightful or all-or-none learning of this task has been consistently positive throughout a number of studies (Trabasso, 1963; Bower and Trabasso, 1963a, b; Trabasso and Bower, 1963). Graphs of backward and Vincentized learning curves (cf. Suppes and Ginsberg, 1963) were uniformly flat at the chance level of 50% correct classifications over trials prior to complete solution. Although the one-step model describes two-category CI data fairly well, it must be extended and modified to account for more complex forms of CI. This paper considers one natural extension, namely, to situations in which the subject must perform two binary discriminations concurrently to solve the task. This is known as the four-category problem and has been studied in detail by several experimenters (Archer, Bourne, and Brown, 1955; Bourne, 1957; Bourne and Haygood, 1959; Bourne and Pendleton, 1959).

In a four-category problem, there are two independent relevant dimensions. For example, if color and form are relevant, then Category 1 may contain all orange circles, Category 2 all orange triangles, Category 3 all blue circles, and Category 4 all blue triangles. If all four categories contain both large and small figures, then size is irrelevant. The question is how one may usefully conceptualize what the subject does in the four-category task in order to develop tractable models of the process? Bourne and Restle (1959) supposed that subjects solve the four-category task by learning two binary subproblems. For the previous example, the two subproblems are defined in terms of the color and form dimensions. The color subproblem is that orange patterns are in classes 1 or 2 and blue patterns are in classes 3 or 4. Likewise, the form subproblem is that circular patterns are in classes 1 or 3 and triangular patterns are in classes 2 or 4.

This notion of subproblem learning supposes that the values of the relevant dimensions become conditioned to the category names reinforced in their presence. This conditioning is shown below.

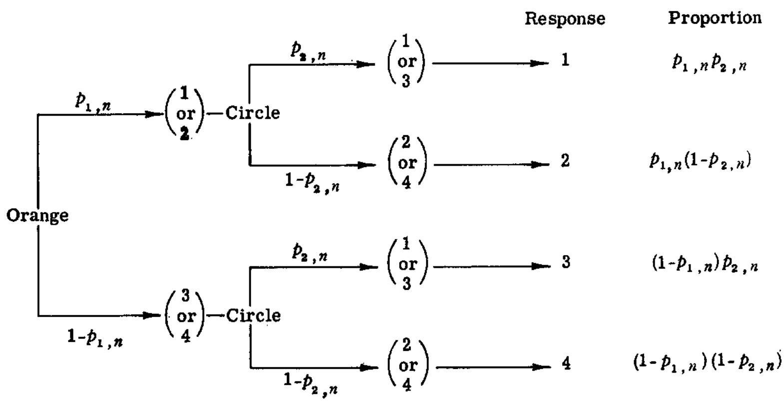
| *Color, Subproblem* | | | *Form, Subproblem* | | |
|---|---|---|---|---|---|
| orange | $<$ | 1<br>2 | circle | $<$ | 1<br>3 |
| blue | $<$ | 3<br>4 | triangle | $<$ | 2<br>4 |

Let $p_{1,n}$ and $p_{2,n}$ represent the probability of a correct response for each subproblem on trial $n$. Using the symmetry of the binary dimensions, the definitions of $p_{1,n}$ and $p_{2,n}$ are

$$p_{1,n} = Pr\,\{1 \text{ or } 2/\text{orange on } n\} = Pr\,\{3 \text{ or } 4/\text{blue on } n\}$$

$$p_{2,n} = Pr\,\{1 \text{ or } 3/\text{circle on } n\} = Pr\,\{2 \text{ or } 4/\text{triangle on } n\}.$$

We now wish to know the probability with which the subject will give any one of the four responses to a particular pattern shown on trial $n$ (e.g., an orange circle). The performance rule is this: the subject generates a pair of covert responses for each relevant attribute and the overtly responds with the common element (intersection) from these two sets. This is illustrated below for the example of an orange circle.




The generation of the two covert response-sets are assumed to be independent.The probability of any response is given by the product of the tree probabilities that lead to it. Because of the independence assumption, it is immaterial in which order the two attributes are considered. The probability of the correct response, 1, in this case is $p_{1,n} \cdot p_{2,n}$. In fact, the model implies that the correct response probability is $p_{1,n} \cdot p_{2,n}$ regardless of the specific stimulus pattern shown on trial $n$.

This, then, is the Bourne-Restle assertion. Its major advantage is an analytic one: it represents the four-category task as the learning of two single dimension subproblems, and it permits direct application of models written for two-category CI learning. For later reference, the Bourne-Restle assumption is written in the form of two equations. Letting $p_{c,n}$ represent the probability of a correct response on the four-category problem on trial $n$:

$$p_{c,n} = (p_{1,n})\,(p_{2,n}) \tag{1}$$

$$p_{c,n} = (p_{1,n})^2 \tag{2}$$

Equation 1 is used when one distinguishes between the two subproblem learning curves. Equation 2 is used when the additional assumption is made that the two subproblems are learned at the same rate, so that $p_{1,n} = p_{2,n}$. Either equation will be referred to generally as the multiplication rule.

Bourne and Restle (1959) made no direct test of Eq. 1, although it was used in generating predictions from Restle's discrimination learning theory (1955) which fit their data fairly well. The purpose of the present experiment was to achieve a direct test of Eq. 1 in the four-category task.

## METHOD

*Subjects.* The subjects were 120 paid students from Foothill Junior College who were randomly assigned to three groups (17 males and 23 females, each) and were run in individual sessions.

*Procedure.* The following instructions were read to all subjects:

> In this problem, we are interested in finding out how college students learn to classify patterns. You will be required to learn a rule or a system which will enable you to classify correctly each card in a set of cards.
>
> This is how we shall proceed. I will show you one card at a time. You will classify the card into one of four categories by calling the card either 1, 2, 3, or 4. At first you must guess the correct classification. After you classify the card, this box will indicate to you the correct classification of the card (illustrated by flashing, in a random order, all four numbers). After a time, you should be able to figure out a rule which will enable you to classify correctly all the cards. The cards are shuffled so that there is no systematic sequence in the presentation.
>
> On each card there will be a pattern consisting of a figure and a line (two cards which differed in all five dimensions were shown). From card to card, the pattern can change in any of five ways so that there are five dimensions to consider. These five dimensions are as follows (pointed out in a random order for each subject: (1) Color: either orange or blue; (2) Shape: either a triangle or a circle; (3) Position of the figure: either on top or bottom of the card; (4) Size of the figure: either large or small; (5) Position of the line: either in the top left or lower right corner.
>
> The solution to the problem will depend upon only two of these five dimensions. By this, I mean that only two dimensions are crucial and only they will determine the category of a card.
>
> Let me illustrate to you what I mean by using two dimensions to classify a card. This example will not contain the dimensions used in your problem but the principle is the same, namely, two dimensions can be combined to yield four unique categories.
>
> Here are four cards labelled 1, 2, 3, and 4, respectively. Note that there are only two dimensions shown on these cards. The figures may be either horizontal or vertical in arrangement and either two or three in number. Thus the dimensions are arrangement and number. If we combine these two aspects by forming all possible combinations, we have four unique classes. Thus, in the example, category one contains two horizontal objects; category two contains two vertical objects; category three contains three horizontal objects; and category four contains four vertical objects. Is that clear? Remember, the example is to illustrate how one can combine two dimensions; the dimensions in the example are not contained in your problem.

Here is a card with the categories and some information concerning your problem. Please read the information. You may keep the card for your reference throughout the problem. O.K. Now let us begin. Here is the first card.

Patterns were presented one at a time on a card holder. The subject paced his verbal responses and the experimenter then showed the correct classification on a lighted panel to the right of the stimulus display. The subject was allowed approximately 4 seconds to view the pattern after reinforcement. A different order was presented to each subject. The order was irregular with the restriction that no two consecutive patterns were of the same category. Cards were rearranged at the end of every 32 trials if the subject had not yet reached the learning criterion of 16 successive correct responses.

*Stimulus materials.* The stimuli were geometric figures drawn in crayon pencil from templates on white 3 × 5 inch file cards. There were 5 binary dimensions: color (orange or blue); form (circle or triangle); size (large or small); position of the figure (top or bottom of card); and position of a line (upper left or lower right corner). Color and form were relevant but independent; all other dimensions were varied independently of each other and were irrelevant. The correct classifications of the patterns were orange: circle-1; orange triangle-2; blue circle-3, and blue triangle-4.

*Experimental groups.* One group learned the four-category problem from scratch; two other groups learned the subproblems. Each subject was given an information card for his reference throughout the experiment. Each card contained the category numbers and the statement: "Remember, the category of a card is determined by *two* dimensions of the pattern." The remaining information on the card differentiated the subjects into three groups:

Group C (Color) was told that the form dimension was relevant. They were given a card which stated that a circle belonged to Categories 1 or 3 and a triangle belonged to Categories 2 or 4. Group C subjects had to learn the *color* half of the problem.

Group F (Form) was told that color was a relevant dimension. They were given a card which stated orange patterns were in Categories 1 or 2 and blue patterns were in Categories 3 or 4. Group F subjects had to learn the *form* half of the problem.

Group CF (Color-Form) learned the entire four-category problem. They were given no advance information concerning the solution other than that two dimensions were relevant.

It should be noted that Group CF had an initial probability of one-fourth of being correct whereas Groups C and F began with an initial probability of one-half of being correct.

## RESULTS

The initial data analyses concern empirical tests of the multiplication rule. The validity of the test follows from three assumptions: (a) that subproblem learning is a fair representation of what the subjects are doing in the four-category task, (b) that rate of learning a subproblem is the same for Group CF as for a "single-cue" Group (C or F) given advance information about the other subproblem, and (c) that performance on the compound problem is determined via the multiplication rule from subproblem probabilities. Later, alternative interpretations of the multiplication rule will be considered. The test does not depend on a particular learning theory, since no assumptions are involved concerning the process by which the subproblems are learned (e.g., incremental or all-or-none). Therefore, if verified, the multiplication rule permits extension of any learning model (for two-category problems) to the more complicated four-category problems.

The mean probabilities of correct responses were obtained in five-trial blocks for each of the three Groups, C, F, and CF. One test of Eq. 1 is how well the points on the CF learning curve are predicted by multiplying the corresponding C and F percentages on each successive trial-block. These results are shown in the middle column of Table 1. To test Eq. 2, C and F subproblems are assumed to be equally difficult and are pooled to obtain an average subproblem learning curve. By squaring the resulting proportion correct in each trial-block, predictions of the CF learning curve are obtained and are shown in the last column of Table 1.

TABLE 1

EMPIRICAL MEAN LEARNING CURVE PREDICTIONS FOR GROUP CF BY SUBPROBLEM GROUPS C AND F

| 5 Trial blocks | Proportion correct reponses | | |
|---|---|---|---|
| | Observed | Eq. 1 | Eq. 2 |
| 1 | 0.300 | 0.316 | 0.316 |
| 2 | 0.430 | 0.458 | 0.462 |
| 3 | 0.455 | 0.448 | 0.442 |
| 4 | 0.545 | 0.530 | 0.490 |
| 5 | 0.665 | 0.638 | 0.640 |
| 6 | 0.700 | 0.684 | 0.684 |
| 7 | 0.760 | 0.804 | 0.789 |
| 8 | 0.840 | 0.810 | 0.810 |
| 9 | 0.900 | 0.878 | 0.874 |
| 10 | 0.945 | 0.920 | 0.922 |
| 11 | 0.940 | 0.920 | 0.922 |
| 12 | 0.945 | 0.950 | 0.951 |
| 13 | 0.965 | 0.985 | 0.986 |
| 14 | 0.955 | 0.990 | 0.990 |
| 15 | 0.975 | 0.995 | 0.996 |
| 16 | 0.990 | 1.000 | 1.000 |
| 17 | 1.000 | 1.000 | 1.000 |
| Mean errors | 18.65 | 18.37 | 18.64 |

The predictions of the CF learning curve in Table 1 appear quite accurate. For instance, during the first ten blocks of trials, the average absolute discrepancy between predicted and observed is 0.023 for Eq. 1 and 0.026 for Eq. 2. The corresponding arithmetic average differences are 0.005 for Eq. 1 and 0.012 for Eq. 2. A good single index of the correspondence is the expected total errors (per subject) which may be obtained by summing the error probabilities over individual trials. The observed average total errors were 18.65, whereas 18.37 and 18.64 are predicted by Eqs. 1 and 2, respectively. Thus, the multiplication rule gives a very accurate prediction of the CF performance.

Similar confirmatory results for the multiplication rule were obtained in a small pilot study carried out prior to the one reported here. In the pilot study, the stimuli were four-letter consonant clusters; the "dimensions" were four pairs of consonants (F, G), (V, W), (X, Y), and (R, S) and one letter from each pair was displayed in the patterns (cf. Bower and Trabasso, 1963a, for a description of the materials). The correct answers in the four-category problem depended upon which members of the two relevant pairs were present in the pattern (e.g., FX $-$ 1, FY $-$ 2, GX $-$ 3, GY $-$ 4). The same experimental design was used: two subproblem groups ($N =$ 10 each) learned one-half of the problem (given advance information about the other half) while a compound group ($N = 20$) learned the four-category problem without prior information. The predictions from Eqs. (1) and (2) of performance of the compound group are shown for 10-trial blocks in Table 2. The average total errors for the compound group was 36.35, whereas 36.93 and 36.79 were predicted by Eqs. (1) and (2), respectively.

TABLE 2

EMPIRICAL MEAN LEARNING CURVE PREDICTIONS OF COMPOUND PROBLEM BY SUBPROBLEM GROUPS (PILOT STUDY)

| 10 Trials blocks | Proportion correct responses | | |
|---|---|---|---|
| | Observed | Eq. 1 | Eq. 2 |
| 1 | 0.280 | 0.235 | 0.235 |
| 2 | 0.390 | 0.403 | 0.403 |
| 3 | 0.480 | 0.442 | 0.442 |
| 4 | 0.545 | 0.516 | 0.518 |
| 5 | 0.625 | 0.583 | 0.585 |
| 6 | 0.690 | 0.701 | 0.706 |
| 7 | 0.770 | 0.798 | 0.801 |
| 8 | 0.775 | 0.864 | 0.865 |
| 9 | 0.890 | 0.874 | 0.874 |
| 10 | 0.955 | 0.921 | 0.922 |
| 11 | 0.965 | 0.970 | 0.970 |
| 12 | 1.000 | 1.000 | 1.000 |
| Mean errors | 36.35 | 36.93 | 36.79 |

These data confirm the assumption of independent subproblem performance, viz., for each subject the generation of the covert response sets corresponding to, say, the cues "orange" and "circle," were independent of one another. Additionally, one may inquire about the independence for individual subjects of the learning rates on the two subproblems. This idea can be expressed in terms of a population of subjects learning the four-category problems. Subject $i$ on trial $n$ has subproblem performance probabilities $p_{i1,n}$ and $p_{i2,n}$. Consider conditionals of the form $Pr(p_{i2,n} = X \mid p_{i1,n} = Y)$; then the assumption of independent learning rates supposes that, for fixed $n$, this conditional probability is equal to $Pr(p_{i2,n} = X)$ and does not depend upon $i$, the subject being considered.

To test this assumption of independent subproblem learning rates across subjects, each CF subject's responses were rescored in terms of his performance on the hypothetical subproblems, and then these derived subproblem performance scores were correlated. The procedure requires elaboration to be understandable. Each response of a CF subject is rescored in two ways: whether it is correct on the hypothetical color subproblem (orange-1 or 2, blue-3 or 4 counted as correct) and whether it is correct on the hypothetical form subproblem (circle-1 or 3, triangle-2 or 4). This rescoring procedure yielded for each subject two separate response sequences over trials.

Imagine that the learning of subproblem $i$ by subject $j$ can be described by the equation

$$p_{ij,n} = 1 - \tfrac{1}{2}(1 - \theta_{ij})^{n-1}$$

We wish to correlate $\theta_{1j}$ and $\theta_{2j}$ which summarize subject $j$'s learning rates on the two subproblems. An estimate of $\theta_{ij}$ is obtained from the total number of errors on subproblem $i$:

$$T_{ij} = \sum_{n=1}^{\infty} (1 - p_{ij,n}) = \frac{0.5}{\theta_{ij}}$$

The estimate of $\theta_{ij}$ is then proportional to the reciprocal of the errors made on subproblem $i$. In case there are no errors, set $\theta_{ij} = 1$.

Using this method, $\theta_{1j}$ and $\theta_{2j}$ (corresponding to color and form subproblems) were estimated for each of the 40 subjects in the CF Group. If the subproblem learning rates are independent, then these pairs of $\theta$ values should be uncorrelated. The product-moment correlation was only $+0.047$, indicating no significant correlation ($z = 0.072$, $p > 0.05$). The corresponding correlation between error scores, $T_{1j}$ and $T_{2j}$, was $+0.14$ and not significant ($z = 0.094$, $p > 0.05$). Performance on one subproblem accounts for practically none of the variance in performance on the other subproblem.

To return to the multiplication rule of Eq. 1, it has been tested by predicting the compound learning curve (for CF) from knowledge of the performance curves for the single-cue groups (C and F). If the subproblem analysis is correct, the multiplication rule should work in the reverse direction, predicting the performance curves of the single-cue groups by the decomposition (into subproblem sequences) of the performance curve for Group CF. Using the subproblem sequences derived from Group CF, the curves predicted in this fashion are shown in Table 3. These are compared with the learning curves obtained from the two single-cue groups (C and F). The last two columns show the results when the subproblems are averaged together, as are the learning curves for the C and F Groups.

TABLE 3

EMPIRICAL MEAN LEARNING CURVE PREDICTIONS OF GROUPS C AND F BY RESCORED GROUP CF

| 5 Trial blocks | Proportion of Correct Responses | | | | | |
|---|---|---|---|---|---|---|
| | Group F | | Group C | | Groups C and F | |
| | Obs. | Pred. | Obs. | Pred. | Obs. | Pred. |
| 1 | 0.575 | 0.565 | 0.550 | 0.545 | 0.562 | 0.555 |
| 2 | 0.740 | 0.752 | 0.620 | 0.605 | 0.680 | 0.662 |
| 3 | 0.730 | 0.660 | 0.600 | 0.655 | 0.665 | 0.658 |
| 4 | 0.780 | 0.710 | 0.680 | 0.690 | 0.730 | 0.700 |
| 5 | 0.850 | 0.805 | 0.750 | 0.785 | 0.800 | 0.790 |
| 6 | 0.855 | 0.845 | 0.800 | 0.800 | 0.827 | 0.822 |
| 7 | 0.935 | 0.855 | 0.860 | 0.860 | 0.888 | 0.858 |
| 8 | 0.915 | 0.920 | 0.885 | 0.880 | 0.900 | 0.900 |
| 9 | 0.950 | 0.960 | 0.925 | 0.930 | 0.938 | 0.945 |
| 10 | 0.980 | 0.980 | 0.940 | 0.950 | 0.960 | 0.965 |
| 11 | 0.985 | 0.995 | 0.935 | 0.945 | 0.960 | 0.970 |
| 12 | 0.985 | 0.990 | 0.965 | 0.960 | 0.975 | 0.975 |
| 13 | 0.995 | 0.990 | 0.990 | 0.970 | 0.992 | 0.980 |
| 14 | 1.000 | 0.985 | 0.990 | 0.955 | 0.995 | 0.970 |
| 15 | 1.000 | 0.995 | 0.995 | 0.975 | 0.998 | 0.988 |
| 16 | 1.000 | 1.000 | 1.000 | 0.990 | 1.000 | 0.995 |
| 17 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| Mean errors | 8.625 | 10.125 | 12.575 | 12.525 | 10.600 | 11.325 |

The subproblem learning curves derived from the CF Group approximate quite well the learning curves obtained from Groups C and F. The average absolute discrepancy between predicted and observed points over the first ten blocks is 0.030 and 0.014 for Groups F and C, respectively; using the pooled data (last columns), the average discrepancy is 0.010. The predictions for Group F are consistently too low for blocks 3-7; as a consequence, the predictions for the pooled data are also slightly too low for these trial blocks. When summated to yield expected total errors, these discrepancies stand out a bit clearer.

## Confusion Matrices

A set of results bearing upon the issue of subproblem learning are the stimulus-response confusion matrices for Group CF. The matrix in Table 4 displays the relative frequencies of the various confusion errors summed over all trials and all subjects in Group CF. If subproblem learning is a correct representation of the behavior, then there should occur some point during training when the subject has learned one subproblem but not the other; at this point, then, the errors that occur will represent confusions only with respect to the unlearned subproblem. It follows that during these trials there should be no occurrences of responses which are erroneous with respect to both subproblems. Thus, the antidiagonal cells are expected to contain the smallest entries in each row of the confusion matrix. This deduction is uniformly supported by the data in Table 4.

TABLE 4

Total Error Confusions for Group CF

| Stimuli | Prop. responses | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | Total |
| Orange circle | — | 0.372 | 0.515 | 0.112 | 196 |
| Orange triangle | 0.297 | — | 0.255 | 0.448 | 192 |
| Blue circle | 0.441 | 0.226 | — | 0.333 | 186 |
| Blue triangle | 0.251 | 0.429 | 0.320 | — | 175 |

Another confusion matrix relevant to the subproblem analysis is that derived prior to learning of the first-learned subproblem. The entries in the matrix of Table 5 represent relative frequencies of confusion errors (summed over trials and subjects) prior to the last error on the first-learned subproblem for the subjects in Group CF. Since, in theory, the matrix is derived from data obtained while performance on both subproblems was at the chance level, the confusion errors should be evenly distributed over the cells in each row of Table 5 (i.e., entries of 0.333). The data in Table 5 follow this rule with the one exception of response 4 to stimulus 1 (but not 1 to stimulus 4, however). A chi-square test could not reject equality of row entries in Table 5 ($\chi^2 = 10.21$, df $= 8$, $p > 0.10$). These data indicate little, if any, gradual learning up to this point in the experiment; conversely, they provide some evidence for all-or-none learning of the first-learned subproblem.

To complete the coverage of results, the confusion matrices for the single-cue groups, C and F, are shown in Table 6. These are summed over subjects and all trials during learning. These data tend to show that the values of the unlearned

TABLE 5

ERROR CONFUSIONS FOR GROUP CF PRIOR TO LEARNING THE FIRST SUBPROBLEM

| Stimuli | Responses | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | Total |
| Orange circle | — | 0.415 | 0.435 | 0.150 | 147 |
| Orange triangle | 0.306 | — | 0.366 | 0.328 | 134 |
| Blue circle | 0.319 | 0.344 | — | 0.336 | 122 |
| Blue triangle | 0.333 | 0.318 | 0.348 | — | 132 |

dimensions (e.g., orange and blue for Group C) are responded to symmetrically, as had been assumed for simplicity in the theory. Thus, for Group C, the numbers 117, 118, 124, and 123 occur in cells whose entries were expected to be equal; likewise, for Group F, the numbers 82, 83, 78, and 95 occur in cells which were expected to be equal. The small entries between 1 and 5 represent failures of subjects to use the prior information which was given to them on a card for continuous reference. For example, although subjects in Group C were informed that circles go into class 1 or 3, there were 3 and 2 occasions, respectively, when an orange circle was put into classes 2 and 4.

TABLE 6

ERROR CONFUSION FREQUENCIES FOR GROUPS C AND F

| Stimuli | Responses | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Group C | | | | Group F | | | |
| | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| Orange circle | — | 3 | 117 | 2 | — | 82 | 2 | 2 |
| Orange triangle | 3 | — | 2 | 118 | 83 | — | 0 | 1 |
| Blue circle | 124 | 0 | — | 5 | 2 | 0 | — | 78 |
| Blue triangle | 1 | 123 | 3 | — | 0 | 1 | 95 | — |

## APPLYING AN ALL-OR-NONE LEARNING MODEL: THE SINGLE-CUE GROUPS

The data will now be analyzed from the vantage point of a particular model regarding the learning process. We begin this analysis with the single-cue groups (C and F), deferring until later the analysis of the CF condition. The model to be used for the single-cue data is a two-state Markov chain developed and employed in a previous paper (Bower and Trabasso, 1963a). In brief, it is supposed that at the beginning of each trial, the subject is either in a presolution state or in a solution state. The presolution state is identified with the state of the subject when he is testing out irrelevant hypotheses which have only a chance correlation with the relevant hypothesis. In general, let $p$ represent the average probability that the subject correctly classifies any randomly chosen pattern while he is in this presolution state. The solution state is entered once the subject begins to attend consistently to the relevant cue; the probability of a correct response when the subject is in this state is unity. The solution state is absorbing; once it is entered, the subject remains there (provided, of course, the problem is not changed). All subjects begin on Trial 1 in the presolution state. On each trial, there is some unconditional probability, $\theta$, that the subject leaves the presolution state and solves the problem.

## Presolution Analyses

The data prior to the last error of each subject in Groups C and F were analyzed according to the expectations of the All-or-None model. The theory supposes that these presolution responses can be represented by a stationary and independent binomial process. Groups C and F were pooled since their respective presolution trials should have the same general characteristics. To test for stationarity, both forward (Suppes and Ginsberg, 1963) and backward learning curves were constructed for trials prior to the last error. These were pooled in five-trial blocks.

Both the forward and backward curves approximated horizontal lines. For the forward curve, the weighted average probability of a correct response was 0.55 and the stationarity test yielded $\chi^2 = 15.78$ ($df = 12$, $p > 0.20$). For the backward curve, the weighted average probability of a success was 0.56 and the $\chi^2$ for stationarity was 4.45 ($df = 10$, $p > 0.90$). Another test for stationarity divided each subject's presolution trials into a first and second half. The average number of errors in the first half was 4.31; in the second half, 4.23 ($t = 0.33$, $df = 33$, $p > 0.05$). These results are consistent with the stationarity implied by the model. The fact that the over-all success rate (0.55) is higher than the a priori one-half is unexplained.

Another prediction of the model is that the sequence of correct and incorrect responses prior to the last error should form an independent series of observations. To test this, the probabilities of a success conditional upon a success or a failure on the preceding trial were compared. There were 1600 transition frequencies for this analysis. The conditional probability of a success was 0.554 following a success and 0.544 following an error on the previous trial. These do not differ significantly ($\chi^2 = 0.17$, $df = 1$, $p > 0.60$).

Consider one final test for the binomial characteristics of the sequence of presolution responses. In a binomial series, the length of a run of successes between adjacent errors should conform to a geometric law. Table 7 displays the frequency distribution of this random variable in the data from Groups C and F. The predictions were derived from the geometric law, and they correspond closely with the empirical distributions.

TABLE 7

DISTRIBUTION OF $H$, THE NUMBER OF SUCCESSES BETWEEN ADJACENT ERRORS FOR GROUPS C AND F

| | $Pr(H = k)$ | |
|---|---|---|
| $k$ | Observed | Predicted |
| 0 | 0.473 | 0.473 |
| 1 | 0.236 | 0.249 |
| 2 | 0.141 | 0.131 |
| 3 | 0.076 | 0.069 |
| 4 | 0.036 | 0.036 |
| 5 | 0.021 | 0.019 |
| 6 | 0.006 | 0.010 |
| 7 | 0.004 | 0.005 |
| 8 | 0.006 | 0.003 |
| 9 | 0.001 | 0.002 |
| 10 | 0.000 | 0.001 |
| 11 | 0.000 | 0.001 |
| 12 | 0.000 | 0.000 |

The results given above do not depend upon estimates of the learning-rate parameter; we next consider a few predictions which do. For this application, Groups C and F were pooled since a likelihood ratio test (Restle, 1961) failed to reject the null hypothesis of equal $\theta$-values ($\chi^2 = 3.84$, $df = 1$, $p > 0.05$). The estimate of $\theta$ was obtained from the average total errors (before learning) via the method of moments. The prediction equations for the All-or-None model are presented elsewhere (Bower, 1961a) and are not duplicated here. Table 8 lists the predictions for the probability distribution of total errors (i.e., proportion of subjects solving after $k$ or fewer errors). The fit is only fair for 40 subjects, but it cannot be rejected by the Kolmogorov-Smirnov one-sample test. The maximum discrepancy (0.094) of predicted from observed proportions occurred at $T = 16$, where no subject scored. The average discrepancy before this point was 0.029.

A list of point predictions derived from the All-or-None model are compared with the pooled C and F data in Table 9. Sequential statistics (runs of errors) are predicted fairly well; standard deviations of errors or trial of last error are over-predicted. The fit generally is not as good as found with simpler cases of single-cue two-category learning (cf. Bower and Trabasso, 1963a).

TABLE 8

DISTRIBUTION OF TOTAL ERRORS, $T$, FOR GROUPS C AND F

| $k$ | $Pr(T \leqslant k)$ Observed | $Pr(T \leqslant k)$ Predicted |
|---|---|---|
| 0-2 | 0.180 | 0.162 |
| 3-5 | 0.387 | 0.354 |
| 6-8 | 0.462 | 0.499 |
| 9-11 | 0.600 | 0.625 |
| 12-14 | 0.675 | 0.722 |
| 15-17 | 0.800 | 0.794 |
| 18-20 | 0.862 | 0.848 |
| 21-23 | 0.925 | 0.887 |
| 24-26 | 0.975 | 0.917 |
| 27-29 | 0.975 | 0.938 |
| 30-32 | 0.988 | 0.952 |
| 33-35 | 1.000 | 0.962 |

Maximum discrepancy = 0.094 at $T = 16$

TABLE 9

SUMMARY STATISTICS OF GROUPS C AND F

| Statistic | Observed | Predicted |
|---|---|---|
| Mean errors | 10.60 | [a] |
| s.d. | 7.92 | 10.59 |
| Mean trial of last error | 22.48 | [a] |
| s.d. | 17.84 | 23.29 |
| Average runs | 6.15 | 6.06 |
| Runs of length 1 | 3.40 | 3.47 |
| 2 | 1.78 | 1.48 |
| 3 | 0.62 | 0.63 |
| Average successes between errors | 1.11 | 1.13 |
| s.d. | 1.33 | 1.55 |

[a] Used to estimate $p$ (0.553) and $\theta$ (0.042).

## Application of the All-or-None Model to the Four-Category Data

The subproblem analysis appears to be a tenable approach to a theory of four-category data. Additionally the subproblem learning of the single-cue groups, C and F, was described fairly well by the All-or-None theory. In our model of four-category learning, we suppose that there are two All-or-None learning processes going on concurrently, one per subproblem, and that their characteristics are the same as revealed in Groups C and F who learned only subproblems. If we do not distinguish between the subproblems, then at the beginning of any trial the subject may be characterized as having already solved either 0, 1, or 2 of the subproblems. Corresponding to each of these levels of knowledge, we indentify 3 states of a Markov chain, all subjects being in state 0 on Trial 1 and eventually becoming absorbed in state 2 when the complete problem is solved. Associated with each state is a probability of a correct response; these are 1/4, 1/2, and 1 for states 0, 1, and 2, respectively. These arise directly from the subproblem analysis, in which the success probability on each subproblem is 1/2 or 1.

Let $\theta$ represent the probability that an unlearned subproblem is learned (solved) in a single trial. It will be assumed that the two subproblems are equally difficult, so that their learning parameters are equal to a common value of $\theta$. It will be further assumed that $\theta$ remains constant over trials in the experiment. In particular, this means that the probability of solving subproblem $A$ per trial is independent of whether subproblem $B$ has already been solved. The trial to trial matrix of transition probabilities for Group CF is given in Eq. 3.

$$
\begin{array}{ll|ccc}
 & & & \text{State on trial } n+1 & \\
 & & 2 & 1 & 0 \\
\hline
 & 2 & 1 & 0 & 0 \\
\text{State on trial } n & 1 & \theta & 1-\theta & 0 \\
 & 0 & \theta^2 & 2\theta(1-\theta) & (1-\theta)^2
\end{array}
\qquad (3)
$$

The descriptive accuracy of this model will be discussed. First, evidence for the 3-state characterization of performance will be considered. Second, we will consider the fit of the model in Eq. 3 to quantitative details of the four-category CF data, in which $\theta$ is estimated from the data of Groups C and F.

## Evidence for Three Performance Levels

The relevant evidence for the intermediate performance level comes from examination of the forward and backward curves prior to the last error in Group CF. The learning curve for an individual is either a single step function from 1/4 to 1 or a two step from 1/4 to 1/2 to 1. The backward learning curve in the former case would be flat at 1/4; in the latter case, it should be at 1/2, then step down to 1/4. The usefulness of the backward curve in identifying the second step (1/2) depends on the proportion of subjects expected to go through that state. This expected proportion is

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

In our case, $\theta$ is small (0.047), so that about 98% of the subjects are expected to go through the intermediate state.

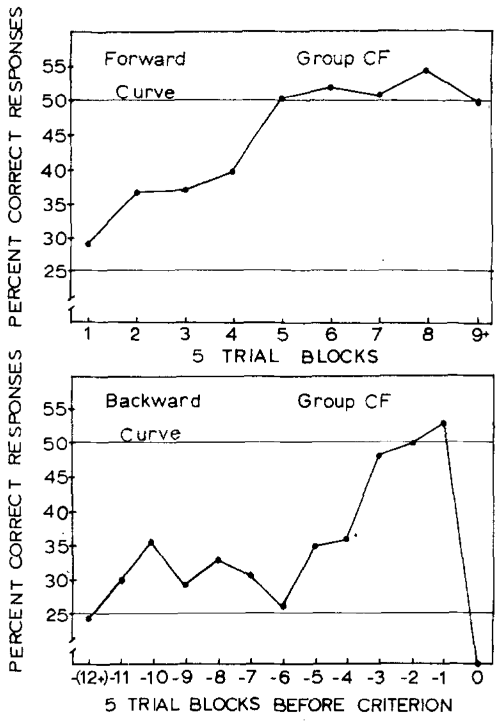



FIG. 1. Forward and backward stationarity curves for Group CF: Proportion of successes prior to the last error.

The forward and backward curves prior to the last error, averaged across subjects in five-trial blocks, are shown in Fig. 1. The top curve in Fig. 1 is calculated by going forward from Trial 1 and dropping each subject as he makes his last error; the bottom curve is calculated by moving backward from the trial of each subject's last error. Both plots reveal the same pattern of results: the average probability of a correct response begins near 1/4 and increases to 1/2 over successive trials prior to the last error. Vincentizing the trials yields the same result. In theory, the transition from 1/4 to

1/2 is a discrete jump for each individual; the gradually increasing curves in Fig. 1 presumably arise because of average across individuals who make the jump at different trials.

PREDICTIONS OF DETAILS OF CF DATA FROM THE MODEL

We now apply the model in Eq. 3 to the CF data, generating predictions which are a function of $\theta$. The mathematical derivations are contained in the Appendix. We could estimate $\theta$ from the total errors for Group CF and then predict other features of CF performance. A more exacting test is to estimate $\theta$ from the data of Groups C and F who learned only one subproblem. This method tests the invariance of the basic parameter as well as the descriptive fit of the model to the CF data.

The parameter estimate $\hat{\theta} = 0.047$ was obtained from the average total errors (10.60) of Groups C and F pooled by using the formula $\hat{\theta} = 0.5/T$. For the 3-state model, the average probability of a correct response on trial $n$ is

$$p_n = 1 - (1 - \theta)^{n-1} + 0.25\,(1 - \theta)^{2(n-1)} \tag{4}$$

This fits the mean learning curve in Table 1 fairly well with $\theta = 0.047$. The mean learning curve for the single-cue subjects is expected to be

$$p'_n = 1 - 0.5\,(1 - \theta)^{n-1}$$

It will be noted that $(p'_n)^2$ is equal to the $p_n$ in Eq. 4; hence, this 3-state model implies the multiplication rule which the data have confirmed.

The expected total number of errors for Group CF is

$$E(T) = \sum_{n=1}^{\infty} (1 - p_n) = \frac{1}{\theta} - \frac{0.25}{\theta(2 - \theta)} \tag{5}$$

Setting $\theta = 0.047$, the prediction is 18.56 mean errors, whereas the observed value was 18.65. For the pilot study using four-letter consonant stimuli, the $\theta$ estimated from the single-cue groups was 0.023. The error for the compound group was 36.35, whereas Eq. 5 predicts 36.28. Thus, the model predicts $E(T)$ accurately using the $\theta$ estimated from results of different subjects learning single-cue subproblems.

In the model, the error distribution is the convolution of two geometric distributions representing the errors committed in states 0 and 1 (see Appendix). The observed and predicted cumulative error distributions for the CF group are presented in Table 10. The fit of predicted to observed values is fairly good. The goodness of fit was evaluated by the Kolmogorov-Smirnov one-sample test. The largest discrepancy (0.135) occurred at the value $T = 21$, but is not large enough to cause rejection of the null hypothesis ($N = 40$ cases, $p > 0.20$).

Finally a list of point predictions of statistics of the CF data are presented in Table 11. There are no constraints imposed upon the comparisons since $\theta$ was esti-

TABLE 10
DISTRIBUTION OF TOTAL ERRORS, $T$, FOR GROUP CF

| | $Pr(T \leqslant k)$ | |
|---|---|---|
| $k$ | Observed | Predicted |
| 0-2 | 0.025 | 0.029 |
| 3-5 | 0.100 | 0.115 |
| 6-8 | 0.225 | 0.228 |
| 9-11 | 0.325 | 0.344 |
| 12-14 | 0.475 | 0.454 |
| 15-17 | 0.500 | 0.551 |
| 18-20 | 0.525 | 0.635 |
| 21-23 | 0.675 | 0.705 |
| 24-26 | 0.775 | 0.763 |
| 27-29 | 0.900 | 0.810 |
| 30-32 | 0.900 | 0.847 |
| 33-35 | 0.925 | 0.877 |
| 36-38 | 0.925 | 0.900 |
| 39-41 | 0.950 | 0.919 |
| 42-44 | 0.975 | 0.933 |
| 45-47 | 0.975 | 0.944 |
| 48-50 | 0.975 | 0.953 |
| 51-52 | 1.000 | 0.959 |

Maximum Discrepancy $= 0.135$ at $T = 21$

TABLE 11
SUMMARY STATISTICS FOR GROUP CF

| Statistic | Observed | Predicted |
|---|---|---|
| Mean errors | 18.65 | 18.56 |
| s.d. | 11.44 | 12.91 |
| Trial of last error | 31.22 | 30.82 |
| s.d. | 17.76 | 23.27 |
| Error runs | 7.65 | 7.67 |
| Length 1 | 3.40 | 3.36 |
| 2 | 1.80 | 1.81 |
| 3 | 1.03 | 1.00 |
| 4 | 0.40 | 0.59 |
| Autocorrelations | | |
| lag 1 | 11.00 | 10.88 |
| lag 2 | 10.62 | 10.46 |
| lag 3 | 9.65 | 10.04 |
| lag 4 | 9.70 | 9.64 |

mated independently from the two single-cue groups. The accuracy of the predictions supports the model and confirms a basic assumption, namely, that the learning parameters for the four-category task can be reconstructed in a direct way from the learning parameter for single-cue problems.

## An Alternative Model for Four-Category Learning

So far, the evidence seems to support a model whose major features consist of the subproblem analysis, all-or-none learning, and the 3-state Markov representation. Call this the subproblem or SP model for later abbreviation. We have considered various alternative models for the four-category task to assess how strongly the evidence favored the SP model as opposed to plausible alternatives. A major desideratum of any alternative model is that it be able to reconstruct the learning parameters for the four-category problem from knowledge of the results of the single-cue learners. It was upon this criterion of parameter-invariance that our alternative models floundered. We were unable to find a successful competitor to the SP model.

It will be informative to outline a particular alternative model in order to illustrate the kind of difficulties encountered. This model is a version of one proposed by Suppes and Ginsberg (1962) for two-category experiments. Each of the four subconcepts (orange or blue circles or triangles) is to be represented by a single stimulus element which is assumed to become conditioned to its correct response in paired-associate fashion. Instances of the same subconcept which differ in irrelevant attributes are considered to be presentations of the same stimulus element. Thus, four-category CI is viewed as similar to the learning of 4 paired-associate items, where the items correspond to the four subconcepts. This will be referred to as the PAL model.

In this model, the learning of the 4 items is assumed to proceed independently in an all-or-none fashion. At the beginning of any trial the subject may have already learned either 0, 1, 2, 3, or 4 of the items. The number of items learned will be identified with the 5 states of a Markov chain. Subjects begin in state 0 and eventually are absorbed in state 4.

If an unlearned item is presented and its correct response reinforced, then with probability $c$ the item is learned on that trial, whereas with probability $1 - c$ it remains unlearned. The transition probabilities derived from these assumptions are the same as those for the 4-element pattern model of Estes (1959). Assume the 4 items are presented with equal frequencies and let $i$ and $j$ index states of the system; then the trial to trial transition probabilities from state $i$ to state $j$ are

$$p_{i,j} = \begin{cases} \left(1 - \frac{i}{4}\right) c & \text{for} \quad j = i + 1 \\ 1 - \left(1 - \frac{i}{4}\right) c & \text{for} \quad j = i \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

The response rules for the system are as follows: to conditioned items, the correct response is given; to unconditioned items, the subject guesses responses from a uniform distribution, being correct with probability 1/4. With these rules, the average probability of a correct response when in state $i$ is $(4 + 3i)/16$. We will now consider some implications of this model and how the evidence bears upon it.

The PAL model implies that an individual's average response probabilities will take on the successive values of 0.25, 0.44, 0.63, 0.81, and then 1.00. The expected numbers of trials that the subject will be in each of the transient error states are $c$, $1.33c$, $2c$, and $4c$, respectively. Since the last error is most likely to occur when the subject is in states 2 or 3, the model implies that approximately the last $6c/8.33c = 72\%$ of the trials prior to the last error should display response probabilities of 0.63-0.81. Turning to the evidence provided by the backward learning curves in Fig. 1, the obtained curves do not have the expected shape. Specifically, both curves in Fig. 1 should rise to between 0.63 and 0.81; instead, they rise to near 0.50 as the SP model predicts.

As a second line of evidence, consider the confusion matrices in Tables 4 and 5. The PAL model, with its assumption of random guessing on unlearned items, predicts that in both Tables 4 and 5 the relative confusion frequencies in each row will be equal to 1/3. Although the equality prediction is well approximated in Table 5, it is not in Table 4. Thus, this evidence goes against the PAL model. On the other hand, the SP model successfully anticipated the difference between the results in Tables 4 and 5.

As a third point, it may be shown that the PAL model does not imply the multiplication rule, for which confirming evidence has been reported. To prove this, expressions for the mean learning curves for the four-category and for the single-cue problems are needed. The PAL model implies that the average probability of a correct response on trial $n$ of the four-category experiment is

$$p_n = 1 - 0.75\,(1 - 0.25c)^{n-1} \tag{7}$$

Similarly, the mean probability for subjects learning the single-cue problem (as in Groups C or F) is expected to be

$$p_n' = 1 - 0.5\,(1 - 0.25c)^{n-1} \tag{8}$$

The multiplication rule would hold for the PAL model in case its predicted $p_n$ is equal to its predicted $(p_n')^2$. Calculations from Eqs. 7 and 8, however, show that $(p_n')^2$ should always be less than $p_n$ if this model is correct. The difference between the two quantities increases then decreases over trials, with a magnitude depending on $c$. For the $c$ value estimated from the single-cue learners (0.188), the maximal discrepancy between $p_n$ and $(p_n')^2$ according to the PAL model should be about 0.063 and should occur around trial 15.

Although the discrepancy is small in magnitude, it is consistent in sign. Thus, the difference $p_n - (p'_n)^2$ should be consistently positive in, say, 5-trial blocks during learning. Inspection of Table 1 shows that this difference is positive in only half of the first 16 blocks of trials. Thus, this evidence goes against the PAL model but agrees with the SP model.

A final point is that the PAL model does not possess the necessary invariance of the learning parameter, $c$. The mean errors for the single-cue groups is expected to be $2/c$. The pooled mean errors for Groups C and F was 10.60, so the estimate of $c$ is 0.188. For the four-category task, the expected errors is $3/c$. Assuming parameter invariance, the prediction for the four-category subjects is 15.90 errors as opposed to 18.65 observed. In comparison, the SP model predicts 18.56 errors in the four-category condition.

In sum, the PAL model is inadequate and there is little need to further flog the dead horse. However, the exercise in flagellation is illustrative of the kinds of explanatory problems encountered by alternative models we have considered.

### Analysis of Derived Subproblem Learning Sequences

At this juncture, the 3-state model appears to give a more accurate account of the CF data. However, discrepancies between data and model do arise when the derived subproblem sequences (from CF subjects) are analyzed according to the expectations of the All-or-None theory. Recall that these derived sequences are obtained by rescoring whether each response was correct with respect to the hypothetical color and form subproblems. The result is two subproblem learning sequences (1's and 0's) for each subject in the CF Group.

In applying the All-or-None model to these derived sequences, it is explicitly assumed that interference or forgetting is minimal. That is, if one subproblem is already learned, then the learning of the second subproblem should never produce a momentary relapse (interference) with the first subproblem performance. This is a stringent assumption in this context since it implies that there is no improvement in performance prior to the trial of the last error on a subproblem. We examine to what extent this identification of the point of subproblem learning is supported by the data.

We first consider the distribution of successes between adjacent errors. The observed distribution is compared in Table 12 with that predicted by the All-or-None model. The fit here is excellent.

Next consider the distribution of the total number of errors. The cumulative distribution for the rescored sequences is shown in Table 13 along with the predictions from the model. The predicted and observed distributions do not significantly by a Kolmogorov-Smirnov one sample test ($p > 0.20$).

In Table 14 are presented some summary statistics for the derives subproblem sequences along with their predicted values. It is to be noted that the fit is quite poor on standard deviations of errors and trial of last error.

TABLE 12

Distribution of $H$, the Number of Successes Between Adjacent Errors for Rescored Group CF

| $k$ | $Pr(H = k)$ Observed | Predicted |
|---|---|---|
| 0 | 0.471 | 0.470 |
| 1 | 0.241 | 0.249 |
| 2 | 0.131 | 0.132 |
| 3 | 0.070 | 0.070 |
| 4 | 0.039 | 0.037 |
| 5 | 0.014 | 0.020 |
| 6 | 0.011 | 0.010 |
| 7 | 0.009 | 0.005 |
| 8 | 0.003 | 0.003 |
| 9 | 0.003 | 0.002 |
| 10 | 0.004 | 0.001 |
| 11 | 0.001 | 0.001 |
| 12 | 0.002 | 0.000 |

TABLE 13

Distribution of Total Errors, $T$, for Rescored Group CF

| $k$ | $Pr(T \leqslant k)$ Observed | Predicted |
|---|---|---|
| 0-2 | 0.212 | 0.195 |
| 3-5 | 0.325 | 0.379 |
| 6-8 | 0.388 | 0.520 |
| 9-11 | 0.585 | 0.628 |
| 12-14 | 0.712 | 0.710 |
| 15-17 | 0.800 | 0.773 |
| 18-20 | 0.850 | 0.821 |
| 21-23 | 0.925 | 0.860 |
| 24-26 | 0.938 | 0.890 |
| 27-29 | 0.950 | 0.924 |
| 30-32 | 0.950 | 0.932 |
| 33-35 | 0.962 | 0.938 |
| 36-38 | 0.975 | 0.944 |
| 39-41 | 1.000 | 0.950 |

Maximum Discrepancy = 0.132 at $T = 8$

TABLE 14

SUMMARY STATISTICS FOR RESCORED GROUP CF

| Statistic | Observed | Predicted |
|---|---|---|
| Mean errors | 11.24 | [a] |
| s.d. | 9.11 | 11.33 |
| Trial of last error | 24.91 | [a] |
| s.d. | 17.72 | 25.60 |
| Error runs | 6.42 | 6.59 |
| Length 1 | 3.72 | 3.86 |
| 2 | 1.59 | 1.60 |
| 3 | 0.74 | 0.66 |
| Average successes between errors | 1.21 | 1.13 |
| s.d. | 1.77 | 1.55 |

[a] Used to estimate $p$ (0.57) and $\theta$ (0.038).

The former sets of analyses gave mildly positive results on fitting the All-or-None model to the derived sequences. We now consider some evidence that is distinctly negative. The model assumes that responses prior to the last error are stationary and independent observations from a binomial series. The data indicate that the derived subproblem sequences are neither stationary nor independent. For the stationarity test, the success probability was compared for the first vs. second half of trials prior to the last error. The percentage successes were 0.542 and 0.607 in the first and second halves, respectively. This is a significant increase in success rate (matched $t = 2.92$, $df = 79$, $p < 0.01$). Along with this increase, successive responses proved to be nonindependent. The conditional probability of a success was 0.592 following a success and 0.544 following an error. The $\chi^2$, based on 1838 observations, is 4.05 ($df = 1$, $p < 0.05$).

Nonstationarity of the derived sequences could arise from several sources, but our data has insufficient power to differentiate among these. One possibility is that retention of the first-learned subproblem may be disturbed or interferred with when the second subproblem is learned. A second possibility is that nonstationarity is produced by some terminal paired-association learning that is necessary to attach the responses to the four pairs of relevent cues. The latter is occasioned by the lack of symmetry in the conventional four-category stimulus-response assignments. For example, if one knows that orange is 1 or 2 and that orange circle is 1, he still requires a further arbitrary association to correctly classify blue circles and triangles (which is 3 and which is 4?) If subjects arrive at this state of knowledge near the end of learning, their average response probability on the form subproblem will be about 0.75 instead of 0.50, as it was prior to learning that orange circle is 1. The possibility of this brief paired-associate end effect has been ignored in this paper in the interests of parsimony (cf. the related discussion in Bower and Trabasso, 1963a); however, it is a good candidate to account for nonstationarity in the derived sequences. Special experiments can be contrived to minimize this paired-associate end effect. These would utilize two-component responses (e.g., A1, A2, B1, B2) that mirror the logical structure of the two relevant stimulus cues to which the responses are assigned.

To summarize our general results, the evidence supports the Bourne-Restle contention that four-category CI learning can be viewed as two concurrent but independent subproblem processes. The evidence in favor of this assertion is (a) the accuracy of the multiplication rule in reconstructing Group CF's performance from the performance curves of the subproblem groups, (b) the symmetry of the stimulus-response confusions in the subproblems and in the compound problem, and (c) the lack of correlation between the learning rates of an individual on the two subproblems. Most of the data from Groups C and F could be accounted for by an All-or-None Markov model. Similarly, most of the details of four-category learning by Group CF were predicted by a 3-state Markov model representing the convolution of two simpler all-or-none learning processes. An important feature of the theoretical analysis was the outcome of parameter-invariance; the learning parameter for the compound problem was reconstructed via the theory from a knowledge of the learning parameters for the single-cue subproblems. Despite the good over-all fit of the model, some deviations from the data were encountered. Future work will aim to understand and resolve these discrepancies.

## APPENDIX

This appendix presents the formulas used to calculate theoretical expressions for the various statistics of the generalized one-element model. The techniques of derivation are not included here and the interested reader is referred to Bower (1961b), Bower and Theois (1963), and Suppes and Atkinson (1960) for these methods.

As in preceding sections, the learning parameter, $\theta$, represents the probability that a subject selects and learns the relevant dimension of a subproblem on a particular trial. Using the matrix given in Eq. 3, where we identify three learning states ($S_0$, $S_1$, or $S_2$) with associated probabilities of correct responses of 1/4, 1/2, and 1, respectively, we have the following:

Average probability of being in state $S_i$ on trial $n$, $w_{i,n}$:

$$w_{0,n} = (1-\theta)^{2(n-1)} \tag{1.1}$$

$$w_{1,n} = 2\,[(1-\theta)^{n-1} - (1-\theta)^{2(n-1)}] \tag{1.2}$$

Average probability of an error on trial $n$, $q_n$:

$$\begin{aligned} q_n &= \tfrac{3}{4} w_{0,n} + \tfrac{1}{2} w_{1,n} \\ &= (1-\theta)^{n-1} - \tfrac{1}{4}(1-\theta)^{2(n-1)} \end{aligned} \tag{1.3}$$

Distribution of total errors, $T$:

(A) Distribution of $t_i$, the number of errors made in transient state $i$:

(1) For $S_0$, let $b_0$ be the probability of no more errors in $S_0$, then

$$b_o = \frac{4\theta(2-\theta)}{3+\theta(2-\theta)}$$

and

$$Pr(t_o = k) = \begin{cases} \tfrac{1}{4} b_0 & \text{for} \quad k = 0 \\ (1 - \tfrac{1}{4} b_o)\, b_o(1-b_o)^{k-1} & \text{for} \quad k \geqslant 1 \end{cases} \tag{1.4}$$

(2) For $S_1$, let $b_1$ be the probability of no more errors in $S_1$, then

$$b_1 = \frac{2\theta}{1+\theta}$$

and

$$Pr(t_1 = k) = \begin{cases} \tfrac{1}{2} b_1 & \text{for} \quad k = 0 \\ (1 - \tfrac{1}{2} b_1)\, b_1(1-b_1)^{k-1} & \text{for} \quad k \geqslant 1 \end{cases} \tag{1.5}$$

(B) Distribution of $T$, a convolution of $t_0$ and $t_1$:

$$Pr\{T = k\} = \begin{cases} \left[\dfrac{\theta}{1+\theta}\right]^2 & \text{for} \quad k = 0 \\ A(1-b_o)^{k-1} + B(1-b_1)^{k-1} & \text{for} \quad k \geqslant 1 \end{cases} \tag{1.6}$$

where

$$A = \frac{b_o(4-b_o)}{4(2-\theta)} \left[\frac{b_1(1-\theta)(b_o-2)}{(b_o-b_1)} + \theta\right]$$

and

$$B = \frac{b_o b_1(1-\theta)(2-b_1)(4-b_1)}{4(2-\theta)(b_o-b_1)}$$

Average total errors:

$$E(T) = \frac{7-4\theta}{4\theta(2-\theta)} \tag{1.7}$$

Distribution of trial of last error, $n'$:

$$Pr(n' = k) = \begin{cases} \left(\frac{\theta}{1+\theta}\right)^2 & \text{for} \quad k = 0 \\ w_{o,n}\, 3 \left(\frac{\theta}{1+\theta}\right)^2 + w_{1,n} \left(\frac{\theta}{1+\theta}\right) & \text{for} \quad k = 1 \end{cases} \tag{1.8}$$

Average trial of last error:

$$E(n') = \frac{3\theta + 2(1+\theta)^2\,[(2-\theta)^2 - 1]}{\theta(1+\theta)^2\,(2-\theta)^2} \tag{1.9}$$

Variance of trial of last error:

$$\mathrm{Var}\,(n') = E(n'^2) - [E(n')]^2$$

where

$$E(n')^2 = \frac{(3\theta - 2)\,[2 - \theta(2-\theta)] + 2(2-\theta)^4}{\theta^2(1+\theta)\,(2-\theta)^3} \tag{1.10}$$

Mean $j$-tuples of errors:

Let $a = \frac{1}{2}(1-\theta)$ and $b = \frac{3}{4}(1-\theta)^2$, then

$$u_j = \frac{(1-\theta)}{\theta(2-\theta)}\, a^{j-1} + \frac{3}{4\theta(2-\theta)}\, b^{j-1} + \frac{3}{4}\frac{(1-\theta)}{(2-\theta)} \left[\frac{a^{j-1} - b^{j-1}}{a-b}\right]$$

Runs

$$R = u_1 - u_2$$

Runs of lenght $i$:

$$r_i = u_i - u_{i+1} + u_{i+2} \qquad \text{(cf. Bush, 1959).}$$

Autocorrelation of errors, lag $k$:

$$c_k = \frac{1}{16\,\theta(2-\theta)}\,[4(5 - 2\theta)\,(1-\theta)^k - 3(1-\theta)^{2k}] \tag{1.11}$$

## REFERENCES

Archer, E. J., Bourne, L. E., Jr., and Brown, G. G. Concept identification as a function of irrelevant information and instructions. *J. exp. Psychol.*, 1955, **49**, 153-164.

Bourne, L. E., Jr. Effects of delay of information feedback and task complexity on the identification of concepts. *J. exp. Psychol.*, 1957, **54**, 201-207.

BOURNE, L. E., Jr., AND HAYGOOD, R. C. The role of stimulus redundancy in concept identification. *J. exp. Psychol.*, 1959, **58**, 232-238.

BOURNE, L. E., Jr., AND PENDLETON, R. B. Concept identification as a function of completenes sand probability of information feedback. *J. exp. Psychol.*, 1959, **56**, 413-420.

BOURNE, L. E., Jr., AND RESTLE, F. Mathematical theory of concept identification. *Psychol. Rev.*, 1959, **66**, 278-296.

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

BOWER, G. Application of the all-or-none conditioning model to the learning of compound responses. Tech. Rep. No. 37, Psychol. Ser., Inst. for Mathematical Studies in the Social Sciences, Stanford University, 1961b.

BOWER, G., AND THEOIS, J. A learning model for discrete performance levels. In R. C. Atkinson (Ed.), *Studies in mathematical psychology*. Vol. 1, Ch. 1. Stanford, Calif. : Stanford Univer. Press, 1963. Pp. 1-31.

BOWER, G., AND TRABASSO, T. Concept identification. In R. C. Atkinson (Ed.), *Studies in mathematical psychology*. Ch. 2. Stanford, Calif.: Stanford Univer. Press, 1963a. Pp. 32-94.

BOWER, G., AND TRABASSO, T. Reversals prior to solution in concept identification, *J. exp. Psychol.*, 1963b, in press.

BUSH, R. R. Sequential properties of linear models. In R. R. Bush and W. K. Estes (Eds.), *Studies in mathematical learning theory*. Stanford, Calif.: Stanford Univer. Press, 1959. Pp. 215-227.

ESTES, W. K. Component and pattern models with Markovian interpretations. In R. R. Bush and W. K. Estes (Eds.), *Studies in mathematical learning theory*. Stanford, Calif.: Stanford Univer. Press, 1959. Pp. 9-52.

RESTLE, F. Statistical methods for a theory of cue learning. *Psychometrika*, 1961. **26**, 291-306.

RESTLE, F. A theory of discrimination learning. *Psychol. Rev.*, 1955, **62**, 11-19.

SUPPES, P., AND ATKINSON, R. C. *Markov learning models for multiperson interactions*. Stanford, Calif.: Stanford Univ. Press, 1960.

SUPPES, P., AND GINSBERG, R. A fundamental property of all-or-none models, binomial distribution of responses prior to conditioning, with application to concept formation in children. *Psychol. Rev.* 1963, **70**, 139-161.

SUPPES, P., AND GINSBERG, R. Application of a stimulus sampling model to children's concept formation with and without overt connection responses, *J. exp. Psychol.*, 1962, **63**, 330-336.

TRABASSO, T. R. Stimulus emphasis and all-or-none learning in concept identification. *J. exp. Psychol.*, 1963, **65**, 398-406.

TRABASSO, T. R., AND BOWER, G. Supplementary report: Presolution reversal and dimensional shifts in concept identification. *J. exp. Psychol.*, 1963, in press.

RECEIVED: June 24, 1963