A mathematical model for the wholist strategy in concept identification*

ROBERT C. HAYGOOD

Arizona State University, Tempe, Ariz. 85281
and

HAROLD D. FISHBEIN and CHARLES F. PINZKA
University of Cincinnati, Cincinnati, Ohio 45221

A model is derived to describe the performance of Ss using the wholist strategy in a concept-identification problem. The general model provides the expected number of trials to solution for all numbers of irrelevant dimensions and numbers of levels within dimensions. Errors to solution are computed for special cases.

Most existing mathematical models of concept identification are based on the assumption that Ss use some form of hypothesis-testing strategy (Levine, 1966; Restle, 1962; Trabasso & Bower, 1968). Despite the prevalence of hypothesis-testing models, there is evidence that some Ss use a rather different approach—the *wholist* strategy first defined by Bruner, Goodnow, & Austin (1956). Although this strategy has been well known for over a decade, there exists as yet no mathematical model for it, perhaps because it is somewhat less interesting mathematically than the hypothesis-testing models. The purpose of this paper is to provide such a model.

DESCRIPTION OF STRATEGY

Like hypothesis testing, the wholist strategy is used in a reception-learning concept-identification problem. In such a problem, the S's task is to identify the correct concept through a process of being shown a series of stimulus patterns, guessing their classifications, and receiving corrective feedback. The concept may contain one or more stimulus attributes; if there are two or more, the statement of the concept must contain an explicit statement of a conceptual rule such as conjunction ("and") or inclusive disjunction ("and/or"). The present discussion applies only to the class of problems whose concepts are defined by unidimensional affirmation or by multidimensional conjunctive rules, e.g., "red and large and square."

In the wholist strategy, the S adopts as his working hypothesis the entire set of attributes in the initial positive instance, i.e., the first instance which contains the designated relevant attribute(s). The rationale for this is that any positive instance necessarily includes all of the relevant attributes, and it is only necessary to eliminate those which are irrelevant to discover the correct concept. The S does this by dropping from his working hypothesis any attributes which are missing from succeeding positive instances; negative instances are ignored. The working hypothesis is reduced in this fashion over successive positive instances until it contains only the relevant attributes, i.e., the S has reached solution.

DEVELOPMENT OF THE MODEL

This development will consider two aspects of the model: a determination of the expected number of trials prior to problem solution; and, for the cases of one and two relevant two-level dimensions, a determination of the expected probability of correctly responding on positive

trials, i.e., the "learning curve" for positive trials. Two procedural assumptions are made for the following development: (1) that there are an equal number of positive and negative trials, and (2) that the initial trial (T_0) is always positive.

Number of Trials

The positive wholist strategy has several interesting properties which are germane to the development of the model. First, the S need not know the number of relevant dimensions to be able to use the strategy successfully. Performance of informed and uninformed Ss is identical. Second, the saliency or vividness of a particular stimulus dimension has no effect on performance, so long as the S knows that the dimension is part of the problem and can discriminate all states or levels of the dimension. Third, there is no inter-S variability in trials to solution; all Ss given the same sequence of stimuli will show identical performance. Thus, the only source of variability in a given condition is the variation in the sequence of stimuli. Finally, the number of trials to solution is independent of the number of relevant dimensions. For example, all other things being equal, it will take the same number of trials to reach solution in a problem with one relevant and five irrelevant dimensions as it will in a problem with two relevant and five irrelevant dimensions. However, the error curves for the two problems may be different.

We assume a set of *positive* trials ranging from 0 to n , i.e., trials on which a positive instance is presented, and a set of relevant (R) and irrelevant (I) dimensions. Each of the I dimensions assumes any one of k equally likely attributes or levels on each trial. The levels of the R dimensions are not equally likely because of the constraint that half of the trials must be positive. In a typical problem, the stimulus dimensions might be the shape, color, and size of figure, with shape the relevant dimension ($R = 1$, $I = 2$) and with each dimension having three attributes. For shape, there are square, circle, and triangle; for color, there are red, white, and blue; and for size, there are small, medium, and large ($k = 3$). With color as the relevant dimension, a positive trial could be defined as a trial in which a red figure appears.

The S's initial working hypothesis includes I plus R attributes, one from each dimension in the problem. On later positive trials, he drops from his working hypothesis any attributes which fail to appear, so his hypothesis following Trial n (the n th positive trial) represents the intersection of sets of attributes appearing in all previous positive instances. It is convenient to view the attributes on T_0 as the initial states of the I plus R dimensions. The problem is solved, and no further errors are made, when each of the I dimensions has changed its initial state at least once on a succeeding positive trial. The probability of change in irrelevant attributes increases when k increases.

After n positive trials (excluding T_0), the probability that each irrelevant dimension will have changed its initial state at least once is

$$\left[1 - \left(\frac{1}{k}\right)^n\right]^I \quad (1)$$

Thus, it can be shown that the probability that all irrelevant dimensions will have changed their initial states by the n th positive trial but not previously is

$$p(n) = \left[1 - \left(\frac{1}{k}\right)^n\right]^I - \left[1 - \left(\frac{1}{k}\right)^{n-1}\right]^I \quad (2)$$

Let the random variable N be the number of the positive trial on which change of initial state of all irrelevant dimensions has first occurred. The expected value of N is

*Preparation of this report was supported in part by Research Grant MH 18249 from the U.S. Public Health Service.

$$\begin{aligned}
E(N) &= \sum_{n=1}^{\infty} n \left\{ \left[1 - \left(\frac{1}{k} \right)^n \right]^I - \left[1 - \left(\frac{1}{k} \right)^{n-1} \right]^I \right\} \\
&= \sum_{n=1}^{\infty} n \left\{ \sum_{i=0}^I (-1)^i \binom{I}{i} \left(\frac{1}{k} \right)^{ni} \right. \\
&\quad \left. - \sum_{i=0}^I (-1)^i \binom{I}{i} \left(\frac{1}{k} \right)^{(n-1)i} \right\} \\
&= \sum_{n=1}^{\infty} n \sum_{i=0}^I (-1)^i \binom{I}{i} \frac{1 - k^i}{k^{ni}} \\
&= \sum_{i=1}^I (-1)^i \binom{I}{i} (1 - k^i) \sum_{n=1}^{\infty} \frac{n}{k^{ni}} \\
&= \sum_{i=1}^I (-1)^i \binom{I}{i} (1 - k^i) \frac{k^i}{(k^i - 1)^2} \\
&= \sum_{i=1}^I (-1)^{i+1} \binom{I}{i} \frac{k^i}{k^i - 1} \\
&= 1 + \sum_{i=1}^I \frac{(-1)^{i+1} \binom{I}{i}}{k^i - 1}
\end{aligned} \tag{3}$$

For example, when $k = 3$ and $I = 5$, we have

$$\begin{aligned}
E(N) &= 1 + \sum_{i=1}^5 \frac{(-1)^{i+1} \binom{5}{i}}{3^i - 1} \\
&= 1 + \frac{5}{2} - \frac{10}{8} + \frac{10}{26} - \frac{5}{80} + \frac{1}{242} \approx 2.576.
\end{aligned}$$

In that there are equal numbers of positive and negative trials, the expected number of negative trials between each pair of positive trials is (negative trials) = $(0)(1/2) + (1)(1/4) + (2)(1/8) + (3)(1/16) + \dots = 1$. Since T_0 is positive and the last trial prior to solution is of necessity positive, by definition of the wholist strategy, the expected number of negative trials will be equal to $E(N)$. Thus, the expected total number of trials (including T_0) prior to solution is equal to $1 + 2E(N)$. Table 1 gives the expected total trials for certain selected values of k and I .

The Learning Curve

Hypothesis-testing models typically assume that the S knows nothing about the correct answer until solution is reached, and thus error performance prior to solution is stationary at chance level. Stationarity will also occur in the wholist strategy if the S chooses to respond indifferently or to select the same response category for every stimulus

prior to solution. While seemingly inefficient, such response strategies free the S's processing capabilities and thus reduce the possibility that he will lose track of the working hypothesis or make some strategic error; furthermore, such response strategies do not in any way delay solution. In either case, the random sequence of stimuli will generate a random sequence of errors at chance level.

In contrast to hypothesis-testing strategies, the wholist strategy provides the S with successively better approximations to the correct solution as the number of attributes in the working hypothesis decreases; presumably this information can be used for correct responding. An efficient response-generating strategy is one similar to that proposed by Restle (1962) in another context. The S compares the set of attributes present on any trial to those remaining in the working hypothesis, i.e., those unchanged attributes in the previous positive trial. If more than half the attributes in the working hypothesis are present on a given trial (positive or negative), he calls the stimulus "positive"; if less than half, he calls the stimulus "negative"; if exactly half, we assume that he "flips a coin" and chooses the two response categories equally often.

The following mathematical development, based on the strategy above, allows one to make predictions about the learning curve. We will first present a formula for the general case, but with $k = 2$, and then discuss in detail the cases where $R = 1$ and 2, and I ranging from 1 to 10. Let $U(x, n, R, I)$ denote the probability that there are x attributes in the working hypothesis on the n^{th} positive trial, with R relevant and I irrelevant dimensions. Further, let $C(x, R)$ be the probability of guessing correctly on a positive trial when there were x attributes in the working hypothesis on the previous positive trial. Then, if $p(n, R, I)$ is the probability of guessing correctly on the n^{th} positive trial, we have

$$p(n, R, I) = \sum_{x=R}^{R+I} U(x, n-1, R, I) C(x, R), \tag{4}$$

where $U(x, n-1, R, I)$

$$= \binom{I}{x-R} \left(\frac{1}{2} \right)^{(x-R)(n-1)} \left[1 - \left(\frac{1}{2} \right)^{n-1} \right]^{R+I-x} \tag{5}$$

and

$$C(x, R) = \begin{cases} \left[\frac{1}{2} \left(\frac{x-R}{2} - R \right) + \sum_{i=\frac{x}{2}-R+1}^{x-R} \binom{x-R}{i} \left(\frac{1}{2} \right)^{x-R} \right] & (x \text{ even}) \\ \sum_{i=\frac{x+1}{2}-R}^{x-R} \binom{x-R}{i} \left(\frac{1}{2} \right)^{x-R} & (x \text{ odd}) \end{cases} \tag{6}$$

It is easily verified that $C(2k, R) = C(2k+1, R)$. For $R = 1$, we find that

$$C(2k, 1) = C(2k+1, 1) = 1/2 + \binom{2k-1}{k} \left(\frac{1}{2} \right)^{2k}$$

For I odd, we have

$$\begin{aligned}
p(n, 1, I) &= U(1, n-1, 1, I) C(1, 1) + U(I+1, n-1, 1, I) C(I+1, 1) \\
&\quad + \sum_{k=1}^{(I-1)/2} [U(2k, n-1, 1, I) \\
&\quad + U(2k+1, n-1, 1, I)] C(2k, 1)
\end{aligned} \tag{7}$$

Table 1 Expected Number of Trials Prior to Solution				
Number of Irrelevant Dimensions (I)	Number of Attributes (k)			
	2	3	4	5
1	5.00	4.00	3.67	3.50
2	6.33	4.75	4.20	3.91
3	7.29	5.33	4.63	4.27
4	8.01	5.78	4.99	4.56
5	8.59	6.15	5.29	4.81
10	10.45	7.33	6.23	5.65
15	11.57	8.04	6.79	6.13
20	12.38	8.55	7.19	6.47

Table 2
Probability of Correct Responding on Successive Positive Trials for R = 1

Positive Trial (n)	Number of Irrelevant Dimensions (I)									
	1	2	3	4	5	6	7	8	9	10
1	.0750	.750	.688	.688	.656	.656	.637	.637	.623	.623
2	.875	.812	.773	.746	.726	.709	.696	.685	.676	.668
3	.938	.891	.854	.826	.803	.784	.768	.754	.742	.732
4	.969	.941	.917	.896	.877	.860	.845	.832	.819	.808
5	.984	.970	.956	.943	.931	.919	.909	.898	.889	.880
6	.992	.985	.977	.970	.963	.957	.950	.944	.938	.932
7	.996	.992	.988	.985	.981	.977	.974	.970	.967	.964
8	.998	.996	.994	.992	.990	.989	.987	.985	.983	.981
9	.999	.998	.997	.996	.995	.994	.993	.992	.991	.990
10	1.000	.999	.999	.998	.998	.997	.997	.996	.996	.995

Table 3
Probability of Correct Responding on Successive Positive Trials for R = 2

Positive Trial (n)	Number of Irrelevant Dimensions									
	1	2	3	4	5	6	7	8	9	10
1	1.000	.0875	.875	.812	.812	.773	.773	.746	.746	.726
2	1.000	.969	.938	.910	.887	.867	.849	.834	.820	.808
3	1.000	.992	.980	.967	.953	.939	.926	.914	.902	.891
4	1.000	.998	.995	.990	.985	.979	.973	.966	.960	.953
5	1.000	1.000	.999	.997	.996	.994	.992	.989	.987	.984
6	1.000	1.000	1.000	.999	.999	.998	.998	.997	.996	.995
7	1.000	1.000	1.000	1.000	1.000	1.000	.999	.999	.999	.999
8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

For I even, we have

$$p(n,1,I) = U(1,n-1,1,I)C(1,1) + \sum_{k=1}^{I/2} [U(2k,n-1,1,I) + U(2k+1,n-1,1,I)]C(2k,1) \quad (8)$$

For example,

$$\begin{aligned} p(n,1,1) &= 1 - \left(\frac{1}{2}\right)^{n+1} \\ p(n,1,2) &= 1 - \left(\frac{1}{2}\right)^n + \left(\frac{1}{2}\right)^{2n} \\ p(n,1,3) &= 1 - 3\left(\frac{1}{2}\right)^{n+1} + 3\left(\frac{1}{2}\right)^{2n} - 5\left(\frac{1}{2}\right)^{3n+1} \\ p(n,1,4) &= 1 - \left(\frac{1}{2}\right)^{n-1} + 3\left(\frac{1}{2}\right)^{2n-1} - 5\left(\frac{1}{2}\right)^{3n-1} \\ &\quad + 7\left(\frac{1}{2}\right)^{4n} \\ p(n,1,5) &= 1 - 5\left(\frac{1}{2}\right)^{n+1} + 5\left(\frac{1}{2}\right)^{2n-1} - 25\left(\frac{1}{2}\right)^{3n} \\ &\quad + 35\left(\frac{1}{2}\right)^{4n} - 21\left(\frac{1}{2}\right)^{5n} \end{aligned}$$

The learning curve for positive trials (after T_0) for I = 1, 2, ..., 10 can be seen in Table 2. This table shows the probability of Ss' responses being correct on the n^{th} positive trial, $n = 1, 2, \dots, 10$.

For R = 2, we find that

$$C(2k,2) = C(2k+1,2) = \frac{1}{2} + \left(\frac{2k-1}{k-1}\right)\left(\frac{1}{2}\right)^{2k-1}$$

For I even,

$$p(n,2,I) = \sum_{k=1}^{I/2} [U(2k,n-1,2,I) + U(2k+1,n-1,2,I)]C(2k,2) + U(I+2,n-1,2,I)C(I+2,I) \quad (9)$$

For I odd,

$$p(n,2,I) = \sum_{k=1}^{I+1/2} [U(2k,n-1,2,I) + U(2k+1,n-1,2,I)]C(2k,2) \quad (10)$$

For example,

$$\begin{aligned} p(1,2,1) &= 1 \\ p(1,2,2) &= 1 - \left(\frac{1}{2}\right)^{2n+1} \\ p(1,2,3) &= 1 - 3(1/2)^{2n+1} + \left(\frac{1}{2}\right)^{3n-1} \\ p(1,2,4) &= 1 - 3\left(\frac{1}{2}\right)^{2n} + \left(\frac{1}{2}\right)^{3n-3} - 7\left(\frac{1}{2}\right)^{4n} \\ p(1,2,5) &= 1 - 5\left(\frac{1}{2}\right)^{2n} + 5\left(\frac{1}{2}\right)^{3n-2} - 35\left(\frac{1}{2}\right)^{4n} \\ &\quad + 3\left(\frac{1}{2}\right)^{5n-3} \end{aligned}$$

The learning curve for positive trials with R = 2 can be seen in Table 3.

The expected number of errors on positive trials after T_0 is shown in Table 4 and is given by

Table 4
Expected Number of Errors on Positive Trials

I	R = 1	R = 2
1	0.500	0.000
2	0.667	0.167
3	0.857	0.214
4	0.962	0.324
5	1.082	0.369
6	1.160	0.450
7	1.248	0.490
8	1.310	0.555
9	1.380	0.590
10	1.432	0.645

$$\sum_{n=1}^{\infty} 1 - p(n, R, I) \quad (11)$$

For example, for $R = 1$, $I = 3$, the expectation is

$$\sum_{n=1}^{\infty} 3\left(\frac{1}{2}\right)^{n+1} - 3\left(\frac{1}{2}\right)^{2n} + 5\left(\frac{1}{2}\right)^{3n+1} = \frac{6}{7} = 0.857.$$

DISCUSSION

With the model in hand, two new facets of the wholist strategy become evident. First, increasing the number of irrelevant dimensions results in an approximately logarithmic increase in difficulty for the wholist S. Second, increasing the number of levels within stimulus dimensions leads to an improvement in wholist performance, provided the irrelevant dimensions are allowed to vary naturally (i.e., positive and negative instances are chosen randomly from the sets of possible positive and negatives).

The question of how many Ss use the wholist strategy, or any similar "shrinkage" or focusing strategy, is controversial. Bruner et al (1956) gave Ss a series of 14 problems and found that 62% of the problems were begun in a wholist manner, i.e., with a hypothesis including all attributes from the first positive instance. This high percentage was challenged by Bourne (1963), who gave each S only three problems and found that only 9% of the problems were begun in a wholist manner. Bourne interpreted the discrepancy as indicating that as Ss become more practiced in the performance required, their increasing sophistication leads to the use of more efficient strategies such as the wholist. Since Bruner et al gave their Ss almost five times as many problems, it would not be surprising that more problems were approached in a wholist manner, presumably the later problems.

The majority of studies of concept identification have used naive Ss serving for a single problem only, and the findings of such studies are commonly inconsistent with any substantial use of the wholist strategy. It is well known, for example, that telling the S the number of relevant dimensions aids performance; that the saliency or obviousness of stimulus dimensions has a pronounced effect, that inter-S variability, even when all Ss in a group are given the identical sequence, is so high that large numbers of Ss are required in most concept-identification experiments; that increasing the number of relevant dimensions usually (though not always) causes a decrement in performance; and that the logarithmic curve for increasing numbers of irrelevant dimensions is contraindicated by the large number of studies showing a linear increase in difficulty (Bourne, 1966).

Thus, on the basis of the evidence, it seems likely that very few naive Ss perform in the manner specified by the present model. Pertinent data for sophisticated Ss are scarce indeed, but suggest that the use of the wholist strategy increases with experience. Further research is clearly needed to establish the course of development of more sophisticated strategies with repeated exposure. The present model and related discussion define what performance should be like when the wholist strategy is in use and provide a possible alternative to interview methods of determining a S's strategy. There is, of course, no question that college students can be instructed to use the wholist strategy for any problem in which the number of stimulus dimensions does not exceed their running short-term memory. For such trained Ss, the model presented here routinely provides a satisfactory prediction of performance. One caution must be added, however. It seems reasonable that high-ability Ss may also acquire a technique for using some of the information from negative instances, and thus show better performance than that predicted by the model. This possibility remains to be explored.

REFERENCES

- BOURNE, L. E., JR. Some factors affecting strategies used in problems of concept formation. *American Journal of Psychology*, 1963, 76, 229-238.
- BOURNE, L. E., JR. *Human conceptual behavior*. Boston: Allyn & Bacon, 1966.
- BRUNER, J. S., GOODNOW, J. J., & AUSTIN, G. A. *A study of thinking*. New York: Wiley, 1956.
- LEVINE, M. Hypothesis behavior by humans during discrimination learning. *Journal of Experimental Psychology*, 1966, 71, 331-338.
- RESTLE, F. The selection of strategies in cue learning. *Psychological Review*, 1962, 69, 329-343.
- TRABASSO, T., & BOWER, G. H. *Attention in learning*. New York: Wiley, 1968.