Transformation rules in concept learning

ROBERT E. LASKY and KENNETH D. KALLIO
University of California at Santa Barbara, Santa Barbara, California 93106

Two major classes of models have been proposed to explain concept learning: strength models and distance models (Hayes-Roth & Hayes-Roth, 1977). The present study demonstrates that subjects abstract transformation rules as defined by the Franks and Bransford (1971) distance model. Transformation rules characterize how the patterns of a concept differ from each other. Transformation rules are inconsistent with strength models, which assume that subjects abstract component features and not relational information characterizing the differences among patterns. Whether a strength model or a distance model is more appropriate in other instances of concept learning is probably a function of task demands, stimulus characteristics, and subject characteristics.

Two major classes of models have been proposed to account for the storage and recognition of similar patterns. Hayes-Roth and Hayes-Roth (1977) have labeled these two classes of models strength and distance models.

According to distance models, subjects construct and store a generic representation of all the defining patterns. The generic representation is typically labeled a prototype. During recognition, patterns are compared with the prototype along some dimension of similarity, or distance metric. All patterns can be ordered with respect to this distance metric, irrespective of whether they are novel or familiar. Recognition is determined by similarity to the prototype and not directly by familiarity. Thus, subjects will "recognize" patterns that are novel but similar to the defining patterns. This is never the case for strength models, which assume that component features of the defining patterns are stored in memory. Subjects store only those feature values that are actually presented. Recognition of a pattern is a function of the number of times that its component features have been presented previously (i.e., their strength). The features of the pattern to be recognized are matched with the stored features in memory. The strength of the stored features determines how confidently the pattern is recognized. Although subjects may "recognize" novel patterns consisting of new combinations of familiar feature values, they should never "recognize" patterns containing novel feature values.

A number of researchers have recently attempted to empirically evaluate the relative merits of strength and distance models. These researchers have attempted to determine whether subjects "recognize" patterns containing feature values that are not representative of

the acquisition patterns. Some researchers (Franks & Bransford, 1971; Neumann, 1977; Posner & Keele, 1968, 1970; Reed, 1972; Goldman & Homa, Note 1) report that subjects do recognize patterns containing unrepresentative feature values if those feature values represent the central tendency of the feature values of the acquisition set. These studies support a distance model explanation. Other researchers (Hayes-Roth & Hayes-Roth, 1977; Neumann, 1974, 1977; Posnansky & Neumann, 1976; Goldman & Homa, Note 1) report that subjects do not recognize patterns containing unrepresentative feature values. These studies support a strength model explanation. Reed (1972) suggests that subjects can perform as predicted by distance or strength models depending on the stimuli, task demands, and subject characteristics.

Neumann (1977) does not conclude from these data that distance models are valid in some situations and strength models in others. Rather, he has proposed a strength model that can explain the performance of subjects in all situations. Neumann labels this model the "interval storage model." It "assumes that the representation of values on continuous dimensions is in the form of intervals rather than points. When discrimination among values is poor, the intervals which represent these values are wide and may overlap other values. Thus, the frequencies with which these values are experienced may increment along with the values actually experienced, resulting in a central tendency" (Neumann, 1977, p. 196).

The interval storage model appears to be a compromise between strength and distance models. Distance models hypothesize that dimensions of similarity are important in recognition. In contrast, strength models hypothesize that subjects store feature values that are not related on a dimension of similarity. The interval storage model hypothesizes that feature values are stored in memory, but the feature values overlap, implying a relationship among adjacent values.

Neumann's (1977) interval storage model can

explain all the data reviewed above. In the present study, we attempt to provide evidence that in some situations subjects do perform as predicted by distance models but not by any strength model, including Neumann's interval storage model. In order to do so, we demonstrate that subjects abstract and store in memory a distance metric, specifically, transformation rules, as proposed by Franks and Bransford (1971). Transformation rules indicate that one pattern differs from another by a specific manipulation or transformation of the component features of those patterns (i.e., by a permutation, addition, deletion, substitution, etc., of the features).

Transformation rules are defined by the relationships among patterns and are not abstracted from individual patterns. The relational nature of transformation rules is inconsistent with strength models since strength models assume that features are abstracted from individual patterns. Clearly, transformation rules cannot be reduced to features.

Subjects were presented sequentially four sets of acquisition patterns. The patterns of each set contained component forms that differed from those in the other sets. For half of the subjects (the control subjects), the same transformation rules described the differences among the patterns in each acquisition set. For the remaining subjects (the interference subjects), the last acquisition set to be presented contained a transformation rule that differed from those in the other acquisition sets. According to Franks and Bransford (1971), subjects should abstract a prototype and transformation rules for each acquisition set presented them. Consequently, the interference subjects may experience interference from the other sets when trying to recognize patterns from the last set. If this occurs, these subjects should recognize patterns that differ from the prototype by a transformation rule that does not characterize the acquisition set to be remembered but does characterize the other acquisition sets presented. According to strength models, such a result should not occur. Forms are the features that are abstracted by subjects from each acquisition set. Since each acquisition set contained different forms, the selective interference outlined above should not occur.

EXPERIMENT 1

Method

Subjects. Eighty-eight undergraduates enrolled in introductory psychology at the University of California at Santa Barbara served as subjects. Participation in this study partially fulfilled an optional course requirement.

Materials. The acquisition stimuli consisted of eight sets of patterns (see Figure 1). Each set contained six different patterns. There were four acquisition sets labeled A that were characterized by the same transformation rules. The remaining four sets, labeled B, were also characterized by a common set of transformation rules. Three transformation rules were common to both Acquisition Sets A and B (a deletion of
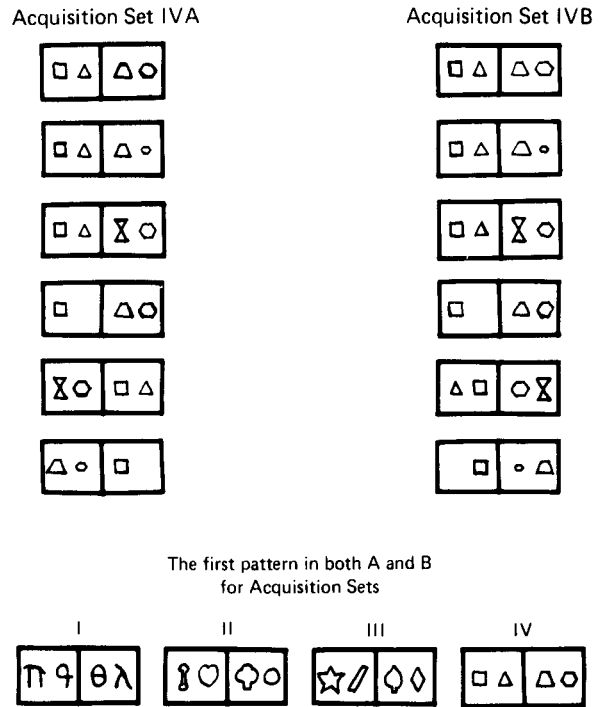

Figure 1. The acquisition sets used in Experiment 1.

one form, a substitution of one form for the other, and a change in the size of one form). A permutation of segments transformation rule (exchanging the two forms in the left half with those in the right half) characterized the differences among the patterns in Set A but not in Set B. A permutation within segments transformation rule (changing the order of forms within each half, or segment, of a pattern) characterized the differences among the patterns in Set B but not in Set A.

The individual patterns consisted of different spatial arrangements of a small set of forms. Four distinct sets of forms were selected. Acquisition Sets IA and IB consisted of the same forms. A different set of forms was used to make Acquisition Sets IIA and IIB. The final two sets of forms were used to make Acquisition Sets III and IV. All the acquisition sets in A differed only in the forms they contained. The same applies to all the acquisition sets in B. By substituting forms from one Acquisition Set A for the corresponding forms from other sets in A, all o the patterns in the acquisition sets labeled A can be generated The same holds for the acquisition sets labeled B.

The acquisition sets were presented to the subjects ir booklets. A booklet contained 12 patterns, the 6 patterns o one acquisition set followed by the same 6 patterns in the same order. The order of the patterns was randomized for each individual subject. The subjects were not informed that each acquisition pattern was presented twice.

The four different recognition patterns are shown ir Figure 2. The transformation distance of each pattern from the
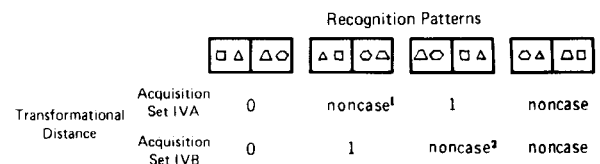

Figure 2. The recognition patterns used in Experiment 1
[1] Test pattern for IVA.   [2] Test pattern for IVB.

prototype is also presented (Franks & Bransford, 1971). According to the Franks and Bransford model, one of these patterns is the prototype of Acquisition Set IVA or IVB, another is a pattern one transformation removed from the prototype, and another is a noncase (i.e., it cannot be derived from the prototype with allowable transformation rules). The final pattern was designated the test pattern and is also one transformation removed from the prototype. It differed from the prototype by a transformation rule that was not characteristic of any of the sets presented to the control subjects. For the control subjects, it was equivalent to a noncase. In contrast, the transformation in question characterized all of the acquisition sets presented to the interference subjects except the set to be recognized (Set IVA or IVB).

The recognition patterns were presented to the subjects in a booklet. Each booklet contained the four patterns in random order followed by the same four patterns in the same order. Each subject was provided a sheet to record his or her recognition ratings. Subjects were required to circle a whole number between −5 and +5 to indicate whether they recognized the pattern or not. A −5 indicated that the subject was positive he or she had not seen the pattern previously. A +5 indicated that he or she was positive that the pattern had been presented previously.

Procedure. The subjects were run in groups. They were given four acquisition booklets and were not to open them until told to do so. The subjects were to look at each pattern for 5 sec. After this presentation, they were to turn the page of the booklet and write four two-digit numbers the experimenter recited on filler sheets between each acquisition pattern. Upon completing the numbers, the subject had 15 sec to draw the pattern before going on to the next page. The experimenter told the subjects when to turn the pages of their booklets. Following the completion of each booklet, subjects were given a 1-min break before going on to the next booklet.

The recognition task was the same for all subjects. They were given a booklet that contained the recognition patterns and a response sheet. They were told how to mark their responses and that they would be allowed 15 sec to view each pattern and to mark their response. The −5 and +5 rating scale that subjects were to use was explained.

Subjects were randomly assigned to one of eight groups with 11 subjects in each group (see Table 1). The control subjects were presented four acquisition sets of the same type (A or B). Two control groups were presented the four acquisition sets in A, and two control groups were presented the four acquisition sets in B. The other four groups were the interference subjects. Two of the interference groups were presented the first three acquisition sets from A and the fourth set from B. The reverse was true for the other two groups of interference subjects. Half of the control and half of the interference groups were tested for recognition 5 min after presenting the acquisition sets. The other four groups were tested for recognition 2 days after presenting the acquisition sets. The hypothesized interference may be more apparent after a 2-day delay, since proactive interference increases with increasing time between acquisition and testing.

Table 1
The Design for Experiment 1

Condition	Acquisition Sets				Delay*
Control	A	A	A	A	0
Control	A	A	A	A	2
Interference	A	A	A	B	0
Interference	A	A	A	B	2
Control	B	B	B	B	0
Control	B	B	B	B	2
Interference	B	B	B	A	0
Interference	B	B	B	A	2

*Delay (in days) between acquisition and recognition.

Results

Table 2 presents the mean recognition ratings for the control and interference subjects. A 2 (delay) by 2 (condition) by 2 (set) by 2 (presentation) by 4 (pattern) analysis of variance was computed for the data. The first three factors were between-subjects factors and the last two within-subjects factors. "Delay" refers to the interval of time between the presentation of the acquisition patterns and the recognition test (either a 0-day or a 2-day delay). "Condition" specifies whether the same transformation rules describe the differences between the patterns in all of the acquisition sets (the control condition) or the set to be recognized differs from the other sets in terms of transformation rules (the interference condition). "Set" refers to which of the two acquisition sets (IVA or IVB) was presented last. "Presentation" refers to the first or the second presentation of the recognition patterns. Finally, "pattern" specifies the four recognition patterns, the prototype, the one-transformation pattern, the test pattern, and the noncase. A significant Condition by Pattern interaction or a Delay by Condition by Pattern interaction would support the Franks and Bransford (1971) distance model.

The Condition by Pattern interaction was significant at the .001 level $[F(3,240) = 7.559]$. Neither condition main effect nor any of the other interactions involving the condition factor was significant. A posteriori contrasts were computed to investigate the significant interaction (Winer, 1971). Subjects in the interference condition rated the test pattern significantly higher at the .01 level than subjects in the control condition. The ratings of subjects in the interference and control conditions did not significantly differ (p > .20)

Table 2
The Ratings of the Recognition Patterns by Subjects in the Control and Interference Conditions

Condition	Recognition Patterns							
	Prototype		One Transformation		Test Pattern		Noncase	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Control	3.90	1.32	2.56	2.17	−2.82	2.76	−3.45	2.11
Interference	3.64	2.09	2.14	2.64	− .64	3.07	−2.61	2.27

Note−The ratings are the average ratings over the two presentations.

for the prototype, the one-transformation pattern, and the noncase. These results are predicted by the Franks and Bransford (1971) model.

The delay [F(1,80) = 4.173], set [F(1,80) = 8.567], and pattern [F(3,80) = 222.054] main effects all proved significant at least at the .05 level. The Set by Pattern [F(3,240) = 9.245], Delay by Pattern [F(3,240) = 3.035], and the Delay by Set by Presentation by Pattern [F(3,240) = 3.035] interactions were also significant at the .05 level. None of these significant effects or interactions is relevant to the hypothesis tested in this study. Consequently, they will not be discussed further.

EXPERIMENT 2

Method

Experiment 2 was conducted to provide a partial replication of Experiment 1. Only one of the four delay by set combinations from Experiment 1 was rerun. Twenty male and 20 female subjects from the same subject pool as in Experiment 1 participated in Experiment 2. All subjects were presented Acquisition Set IVB as the final acquisition set. All subjects were presented the recognition patterns after a 2-day delay. Half of the subjects were in the control condition (10 males and 10 females); the other half were in the interference condition. The methodology was identical in all respects to that used in Experiment 1.

Results

A 2 (condition) by 2 (sex) by 2 (presentation) by 4 (pattern) analysis of variance was computed for the data. "Sex" specifies whether the subjects were males or females. All other factors correspond to those described in Experiment 1. The Condition by Pattern interaction was significant at the .05 level [F(3,108) = 2.733], replicating the results of Experiment 1. Again, subjects in the interference condition rated the test pattern significantly higher than the subjects in the control condition (see Table 3) as hypothesized by the Franks and Bransford (1971) distance model. The sex [F(1,36) = 7.176], pattern [F(3,36) = 156.951], and presentation by pattern [F(3,108) = 2.883] effects were also significant at the .05 level. They will not be discussed further, since they do not bear on the hypothesis investigated.

GENERAL DISCUSSION

The results of Experiments 1 and 2 are predicted by the Franks and Bransford (1971) distance model. However, it is not necessary to conclude from these experiments that transformation rules are used to compute recognition ratings in the manner hypothesized by Franks and Bransford. The information that subjects abstract concerning how the acquisition patterns differ may merely define boundary conditions for instances of the concept defined by those patterns. Subjects may learn that certain variations can occur among a set of patterns while other variations cannot. However, it is clear from Experiments 1 and 2 that subjects do abstract generalizations that characterize the differences among the defining patterns. This is information that cannot be defined by component features. Therefore, it is inconsistent with strength models.

Although the need to assume that subjects abstract transformation rules in concept learning tasks is disputed, the fact that subjects abstract these kinds of rules in other tasks is well established. Indeed, the design and rationale for the present study are similar to those of a study by Reber (1969) demonstrating that subjects learn some artificial languages by abstracting the syntactical structure of the sentences rather than the specific combinations of symbols defining those sentences. Much of the research demonstrating that subjects abstract rules that cannot be reduced to terminal elements (whether they are words, symbols, or features) has been stimulated by Chomsky's (1957, 1965) linguistic theory. The Franks and Bransford (1971) distance model is clearly influenced by Chomsky's theorizing. Their model can be described as a formal grammar consisting of terminal elements (component features), a start symbol (the prototype), and a number of production rules (transformations) specifying how the terminal strings (patterns) of that language (concept) are to be generated (Hopcroft & Ullman, 1969).

Chomsky and his followers are interested in describing those terminal strings that are well formed or grammatical and those that are not. Franks and Bransford share this interest (i.e., differentiating cases from noncases), but in addition, they want their theory to define the similarity between an allowable pattern and the prototype. Thus, they have equated the number of transformation rules involved in generating a pattern from the prototype with the similiarity or distance between the pattern and the prototype. Similarity to the prototype determines the confidence with which the

Table 3
The Ratings of the Recognition Patterns by Subjects in the Control and Interference Conditions in Experiment 2

Condition	Prototype		One Transformation		Test Pattern		Noncase	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Control	4.08	1.18	3.78	1.07	−3.15	1.91	−2.70	2.38
Interference	3.83	1.36	3.00	2.01	−1.73	4.15	−3.03	1.94

Note—The ratings are the average ratings over the two presentations.

pattern is recognized. Their model can be characterized as a derivational theory of recognition.

Hayes-Roth and Hayes-Roth (1977), Neumann (1977), Posnansky and Neumann (1976), Reed (1972), and Goldman and Homa (Note 1) demonstrate convincingly that in some situations subjects perform as predicted by strength models and not by distance models. Other studies (Neumann, 1977; Posner & Keele, 1968, 1970; Reed, 1972; Goldman & Homa, Note 1; the present study) indicate that in other situations distance models provide a more appropriate description of their behavior. These studies support Reed's (1972) hypothesis that subjects learn concepts in a manner predicted by both strength and distance models, depending on the specific situation. When the similarity among patterns is emphasized, subjects may abstract prototypes and distance metrics. When the defining patterns differ by highly discriminable feature values, subjects may abstract and store the frequencies of those feature values.

The major difficulty with Reed's (1972) two-model hypothesis concerns its implication for uncued recognition. Reed's two-model hypothesis implies that the subject knows "how to recognize" patterns before the patterns are recognized. Specifically, a subject must know whether features should be abstracted from the pattern to be recognized and compared with stored features or whether the distance between the pattern and prototype is to be computed. This implication seems unimportant in the present study, since the subject is told the class of patterns he or she is trying to recognize. Presumably, the subject has the generic representation in mind when the recognition patterns are presented. However, this implication becomes critical in situations in which the subject is not cued as to the class of patterns to be recognized. We can certainly recognize a pattern without being warned as to the kind of pattern it is. If we recognize all things in the same manner, there is no problem. If, however, we wish to argue that there is more than one way to recognize a pattern, then the problems associated with selecting the appropriate recognition process must be confronted.

One solution hypothesizes that recognition involves a two-stage process, an initial stage when the generic representation is tentatively identified, and a second stage when the recognition judgment is finalized. Identification of the generic representation during the initial stage may be accomplished by the same process for all patterns. However, once the tentative generic representation is identified, recognition proceeds in the manner dictated by the stored concept. The generic representation may be a multidimensional mode (strength model) or a multidimensional mean (distance

model), depending on the characteristics of the defining patterns, the acquisition task, and subject characteristics. The recognition judgment will also vary depending on these factors. If the similarity among the defining patterns is perceived and stored (a distance model), the recognition comparison is defined by the relationship between the generic representation and the patterns to be recognized on the same dimensions of variation that describe the similarity among the defining patterns. If the feature values are processed as unrelated features (a strength model), the recognition comparison merely specifies whether or not the component feature values of the pattern to be recognized match the stored feature values. Regardless of the validity of this model, the present study demonstrates that in some tasks subjects abstract information that is inconsistent with the kind of information specified by strength models.

REFERENCE NOTE

1. Goldman, D., & Homa, D. *Prototype abstraction as a function of discriminability and number of instances defining the prototype.* Paper presented at the Southeastern Psychological Association, Hollywood, Florida, May 1974.

REFERENCES

CHOMSKY, N. *Syntactic structures.* The Hague: Mouton, 1957.

CHOMSKY, N. *Aspects of the theory of syntax.* Cambridge, Mass: The MIT Press, 1965.

FRANKS, J. J., & BRANSFORD, J. D. Abstraction of visual patterns. *Journal of Experimental Psychology,* 1971, 90, 65-74.

HAYES-ROTH, B., & HAYES-ROTH, F. Concept learning and the recognition and classification of exemplars. *Journal of Verbal Learning and Verbal Behavior,* 1977, 16, 321-338.

HOPCROFT, J., & ULLMAN, J. *Formal languages and relation to automata.* Reading, Mass: Addison-Wesley Publishing Company, 1969.

NEUMANN, P. G. An attribute frequency model for the abstraction of prototypes. *Memory & Cognition,* 1974, 2, 241-248.

NEUMANN, P. G. Visual prototype formation with discontinuous representation of dimensions of variability. *Memory & Cognition,* 1977, 5, 187-197.

POSNANSKY, C. J., & NEUMANN, P. G. The abstraction of visual prototypes by children. *Journal of Experimental Child Psychology,* 1976, 21, 367-379.

POSNER, M., & KEELE, S. On the genesis of abstract ideas. *Journal of Experimental Psychology,* 1968, 77, 353-363.

POSNER, M., & KEELE, S. Retention of abstract ideas. *Journal of Experimental Psychology,* 1970, 83, 304-308.

REBER, A. Transfer of syntactic structure in synthetic languages. *Journal of Experimental Psychology,* 1969, 81, 115-119.

REED, S. Pattern recognition and categorization. *Cognitive Psychology,* 1972, 3, 382-407.

WINER, B. J. *Statistical principles in experimental design.* New York: McGraw-Hill, 1971.