Cognitive Psychology **39,** 69–115 (1999)
Article ID cogp.1999.0718, available online at http://www.idealibrary.com on IDEAL®

Concepts and Transformational Knowledge

Safa R. Zaki and Donald Homa

Arizona State University

The effect of exposure to principled change in concept formation was investigated in four experiments. In Experiment 1, participants were trained on either patterns that transformed systematically or control patterns that were distorted randomly. Training on transformational patterns produced concepts that were more resistant to false intrusions and decay. Experiment 2 separated the relative influences of transformational knowledge and pairwise similarity. Participants were able to identify the next pattern in a transformational sequence even though the foils were closer to the training patterns. Experiment 3 investigated whether participants use transformational information in a speeded categorization task. Participants were faster at classifying patterns that continued a transformational path than patterns that fell off the path, only if they had trained on the transformational patterns in a systematic order. Experiment 4 used multidimensional scaling to explore the psychological structure of transformational knowledge following training. Analyses revealed clear evidence of a transformational path with systematic training. Implications for theories of similarity and categorization are discussed. © 1999 Academic Press

Objects in the natural environment evolve in form. Things change their shape and appearance as they grow. Weathering of the environment produces systematic changes in our landscapes. The seasons result in predictable sequential alterations in plant life. Not only is change a constant in our environment, change is often constrained in its magnitude and direction. In other words, change is often principled. The aim of this research is to explore the effect of exposure to this systematic change on categorization behavior.

We believe that knowledge of how objects change is a largely overlooked but fundamental component of conceptual knowledge. Consider for a mo-

This research was based on a dissertation submitted by the first author to Arizona State University and was supported by the Arizona State University Regents Dissertation Award and a Departmental Dissertation Award to the first author. The first author would like to express her gratitude to her committee members, Steve Goldinger, Todd Maddox, Susan Sommerville, and Guy Van Orden. We would also like to thank Dorrit Billman, Keith Holyoak, John Kruschke, Robert Nosofsky, Mark Steyvers, Huff Templeton III, and two anonymous reviewers for helpful comments on earlier drafts of the manuscript.

Please address correspondence to the first author, who is now at the Cognitive Science Program, Indiana University, Psychology Building 376B, Bloomington, IN 47405. E-mail: szaki@indiana.edu.

0010-0285/99 $30.00
Copyright © 1999 by Academic Press
All rights of reproduction in any form reserved.

ment the changes that a tadpole undergoes before it reaches the final stages of becoming a frog. If one were to look at these entities separately, having no knowledge of the nature of these changes, one might find it difficult to classify them as belonging to the same category. However, given the intermediate steps between the tadpole and the frog, it becomes easier to identify the two examples as being different forms of the same category.

For years, researchers have studied systematic changes in a variety of domains. For example, some researchers have argued that facial changes that occur with aging closely follow the precise mathematical specification of two factors (Cutting, 1978; Pittenger & Shaw, 1975). Biologists have expended great effort in tracking the progression of growth and in recent years a variety of biological growth models have received considerable attention (Grierer & Meinhardt, 1972; Meakin, 1986). Indeed, the branch of biology called morphogenesis (Phillips, 1975) deals explicitly with the epigenetic influences that are responsible for the development of an organism, noting that some characteristics of the adult unfold at various times during its development (Ebeling, 1991). Often, the diagnosis of disease depends on the progression of the disease as opposed to the symptoms at any one time, as in the bifurcation of cellular development of normal and cancerous cells (Muller, Rambaek, Hovig, & Hovig, 1991).

The utility of understanding transformations appears to be fundamental, yet only a handful of studies have addressed the cognitive effect of processing these transformations. Alexander and Enns (1988) used transformations to explore developmental changes in category boundaries, using exemplars that spanned a continuum separating two distinct categories. Findings suggested that category boundaries became less fuzzy with age. DeRosa and Tkacz (1976) demonstrated that memory search was speeded if the memory set was composed of sequenced, rather than nonsequenced, elements. Hull (1920) may be said to have used quasitransformations, demonstrating that concepts were learned more rapidly when learned in the order from simple to complex. In the face perception literature, considerable attention has been given to the transformation of a face as it ages. For example, Pittenger and Shaw (1975) demonstrated that participants were highly sensitive to changes in a cardioidal strain of a profile of a face. The degree of this strain, which is a nonlinear topological transformation, was correlated with the relative age rating of a face. Also, Pittenger, Shaw, and Mark (1979) showed that participants perceived this strain as growth even when it was applied to inanimate objects, suggesting that subjects have abstract knowledge of the growth transformation.

Certainly, researchers have explored the ability to extrapolate (Brehmer, 1974; Busemeyer, Byun, DeLosh, & McDaniel, 1997; DeLosh, Busemeyer, & McDaniel, 1997; Koh & Meyer, 1991) beyond the training items seen. However, the literature base lacks studies that investigate how exposure to principled change affects conceptual knowledge when the changes are identified as transformations of the same object.

Researchers in perception have also explored the importance of transformations on a percept. For example, studies in representational momentum (e.g., Finke & Freyd, 1985; Freyd, 1987; Pinker, Choate, & Finke, 1984) suggest that the representation of an object in motion includes both actual and implied states. When a pattern is seen to change its orientations so that rotation is implied, participants are perceptually biased to extend the path of an object undergoing temporal change (e.g., Freyd & Johnson, 1987) and these implied states are determined by the global path, rather than by local changes, of a stimulus (Verfaillie & D'Ydewalle, 1991). However, this line of research has focused on the memory and momentum of the representation as opposed to the knowledge of the transformation.

Finally, theories in object perception have explored whether depth rotated objects may be represented by viewer invariant (Biederman & Gerhardstein, 1993) or viewer sensitive information, (Lawson & Humphreys, 1996). Recently others have suggested that the spatiotemporal character of visual stimuli is used in object recognition (Stone, 1997; Ullman, 1979). For example, Stone showed that when the order of images in a learned sequence was reversed, the ability to recognize the object decreased significantly. Stone suggests that the ability to utilize simple spatiotemporal sequences may be deeply imbedded in the biological visual system. These researchers view static object recognition as a special case of the ability to recognize objects from sequences of their motion.

An apparent exception to the paucity of research on transformations and their affect on category performance is research conducted by Rips (1989). In his experiments, Rips describes transformations to his participants through the telling of a story. These stories would, for example, describe an imaginary animal undergoing certain transformations that caused many of its surface or essential properties to resemble those of another animal. These changes were either a result of maturation (essence condition) or the result of some catastrophe (accident condition). Participants then rated the typicality, similarity, and likelihood of category membership of the resultant animal. His findings indicated that participants would sometimes rate the stimulus as being more similar to one category but more likely to belong to another. If the changes were a result of maturation, then the item would be rated as a member of the new category but similar to the old. If the changes were a result of an accident then the participants would show the reverse effect.

Therefore, although Rips manipulated transformations, the focus of his study was more on the essential versus accidental properties and whether similarity and categorization can be dissociated, rather than on the results of exposure to transformations. In addition, his participants were exposed to the transformations by way of a story but were not given visual illustrations of these changes.

CATEGORIES, COHERENCE, AND CAUSALITY

Our assumption is that transformational knowledge contributes to the coherence of a concept; a subject, once exposed to a succession of changes, may make the inference that these various states cohere into the definition of a single object, one that has the potential for dynamic change. This approach may be contrasted with research that has similarly explored the question of what makes a concept coherent, or, alternatively, what provides the glue that binds the members of a category together.

One approach, following failures to observe ‘‘common features’’ of members of a category (e.g., Rosch, 1975b; Hampton, 1987), has been to posit perceived causality, naïve theories, and/or goals as the basis for the coherence of members of a category. For example, Medin, Wattenmaker, and Hampson (1987) suggest that an underlying concept or cause helps to structure what would otherwise be independent properties within a category. Murphy and Medin (1985) argue that concepts are coherent to the extent that they fit peoples' background knowledge about things. This view asserts that conceptual coherence is important but is derived from influences that are external to the stimulus set (e.g., Murphy & Medin, 1985; Rips & Collins, 1993; Medin et al., 1987) viz., two members of a category are similar to the extent that they fit one's personal theory or goal.

There is little question that category coherence is an inference that arises early in the development of concepts. Demonstrations of successful distinctions between inanimate and animate objects by children as young as 3 years of age (e.g., Keil, 1989), primitive reasoning about creation and beginning (Gelman & Kremer, 1991), and even support for scientific rationality in childhood (Samarapungavan, 1992) suggest that inferences about what makes concepts coherent begins at an early age and that internal mechanisms are responsible for this outcome.

We suggest that transformational knowledge provides one fundamental type of conceptual coherence. Exposure to the successive changes of an object facilitates the gluing of those states into a single concept. Our view of transformational knowledge, with its emphasis on successive change from a common origin, bears a parallel to the Aristotelian views of categories.

> Nature means . . . the generation of ‘‘growing’’ things. It means also an inherent something out of which a thing begins to grow . . . that in natural beings there inheres a source of their motion, things are said to ‘‘grow’’ when they increase . . . as in the case of embryos. . .What all beginnings have in common is that they are points of departure either for being, or becoming, or knowing. (all passages from The Metaphysics, book Delta, translated by Hope, 1968)

TRANSFORMATIONAL KNOWLEDGE AND MODELS OF CATEGORIZATION

Research in categorization theory has concentrated on two broad lines of inquiry, the identification of those variables thought to shape concepts, and the development of quantitative models that capture categorization performance. Variables such as category size, distortion level, and number of categories have been found to be critical in determining performance (Homa, 1984). For example, greater category size has been shown to increase correct classification in transfer tests, especially when coupled with increased distortions of the patterns from their central tendency (Homa & Vosburgh, 1976). Instance frequency is important to subsequent classification (Nosofsky, 1988), although its importance may be restricted to low and intermediate levels of learning, especially for small category sizes (Homa, Dunbar, & Nohre, 1991).

The proliferation of formal quantitative models, especially within the past 10 years, has largely evolved around whether categorical representations are dominated by specific exemplars (e.g., Nosofsky, 1986, 1988; Shin & Nosofsky, 1992), abstractions (e.g., Knapp & Anderson, 1984; Posner & Keele, 1968, 1970), features (e.g., Estes, 1986; Hintzman, 1986), or decision bounds (Ashby & Maddox, 1994b; Maddox & Ashby, 1993). Another relatively recent advance has been the implementation of neural network/adaptive models (e.g., Kruschke, 1992) to categorization data, using back-propagation and hidden nodes as a means of adjusting performance during learning. None of these models, however, has addressed the utility of transformational knowledge.

Although the focus of this paper is not on the models of categorization, but on a phenomenon of human cognition that has gone for the most part unstudied, it is important to consider how the various models might account for a demonstration of enhanced categorization performance as a function of exposure to principled change. Consider three general cases that capture the essence of some of the conditions explored in the following experiments. In condition 1, the participant observes a stimulus that undergoes systematic change from initial to final form; in condition 2, the participant observes the same stimuli but in a random order; in condition 3, the participant observes stimuli that are matched in terms of distance from the prototype but which have no systematic relationship to each other. Exemplar models of classification (e.g., Medin & Schaffer, 1978; Nosofsky, 1988) must predict that the transformational conditions (conditions 1 and 2) would produce identical transfer performance because the learning and transfer stimuli are identical, with the learning conditions differing only in their order of presentation. Specifically, the classification algorithm for the Generalized Context Model (Nosofsky, 1988) contains only interexemplar similarities, exemplar frequencies, and sometimes bias parameters. Nonetheless, these models would probably predict an overall advantage of the transformational conditions (1 and 2) relative to the random condition (3) since the transformational conditions are likely to have more extreme within-category, interitem similarities than the random condition.

Feature models (e.g., Gluck & Bower, 1988; Hayes-Roth & Hayes-Roth, 1977) appear to make predictions similar to the exemplar models since the transformational conditions would have identical feature sets and therefore similar transfer predictions. Again, the feature sets (including higher order featural configurations) for the transformational conditions might be packed with more extreme feature similarities than the random condition, thereby favoring transformational performance over a random condition.

A pure prototype model, in which only similarities of the transfer instances to the prototypes are critical to transfer, predicts no difference between transformational and random presentations since the exemplar distances to the prototype (but not among the patterns) are identical in all conditions. If the participant abstracts not only the prototype (mean) and breadth (variance) of the category (Homa & Vosburgh, 1976) but a vector for distortions, then the participant could, conceivably, generate distortions that are systematically removed from the prototype. However, even this model would not predict differences among the transformational conditions.

Decision-bound models postulate that participants learn to assign responses to different regions of the perceptual space. When categorizing an object, the participant determines the region in which the percept falls and appropriately emits the response. The bound is the region between these two response regions. Exemplar information is not needed for classification decisions but is necessary to construct the bound. Predictions of how decision-bound models might fare are difficult without knowing the exact or even approximate location of the bound. However, there are no a priori reasons to predict differences in performance as a function of the order of presentation in learning.

Certainly, a generative model, in which the participant abstracts a rule for how patterns can systematically change, would predict differences among the transformational conditions. In effect, the participant must have knowledge of more than the central tendency, variance, and/or specific instances should differences emerge among the transformational conditions.

EXPERIMENT 1

Experiment 1 investigated how the knowledge of the way in which an entity can change may aid future classification and recognition. It is useful at this point to operationalize what we mean by a transformation. In this and the following experiments, we define it to be the successive change that interpolates the states between two endpoints. In these experiments, we used nine-dot patterns where the transformations were interpolated states between the prototype and some high level distortions. Anecdotal evidence from informal piloting of these stimuli produced the impression from many subjects, at least following systematic exposure, of an unfolding of an event and that even the highest levels of distortion were now readily comprehended and seen to have a common origin. As it is operationalized in these experiments, a transformation is assumed to be linear. This is just one possible kind of transformation that we have chosen to investigate in the article, but clearly other types of transformations exist and are processed by people. For example, the changes that occur in a face over time are not always linear (Cutting, 1978; Pittenger & Shaw, 1975).

The main goals of the first experiment were as follows. We wanted to compare conditions that received training on patterns that were distorted from some initial form to a final form to conditions in which the patterns were equally distorted but where no guiding principle governed successive distortions. In addition, we wanted to investigate the effect of viewing these transformations in a systematic or scrambled order.[1]

We trained subjects on three categories of patterns that were either randomly distorted patterns or stages in the transformation of the prototype into a high level distortion. Subjects were explicitly told that items within a category were different manifestations of the same object. In the transformational conditions, each category had multiple paths or vectors of transformations. This was motivated by the concern that transformations of real world objects may manifest themselves along several possible paths. In order to assess the affects of order, subjects either saw these patterns scrambled within a category, scrambled within a transformational path, or in an ordered fashion within a transformational path.

Subjects in the random conditions were trained on patterns that were yoked to be at a distance[2] from the prototype equal to a corresponding pattern in the transformational conditions. However, the fundamental difference between the two types of conditions was that these distortions were not systematic in nature. Because pilot studies suggested that subjects in the random conditions took longer to learn the category, we decided to run two random conditions. In one random condition, all participants received a specified number of learning blocks. In another random condition, participants received a variable number of learning blocks so that their terminal learning performance would approximately match that of the participants in the transformational conditions.

After training was completed, subjects received a transfer test that included a variety of old and new patterns. The new patterns included the

[1] We use the word ‘‘scrambled’’ rather than the word ‘‘random’’ to refer to a random order of presentation. This wording was chosen so as to avoid confusion between random distortions of stimuli and random ordering of stimuli.

[2] We use the word ‘‘distance’’ to refer to Euclidean distance in this paper.

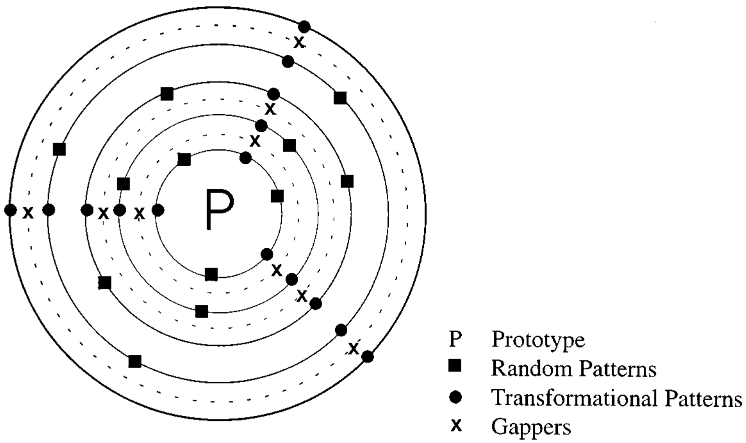

FIG. 1. A schematic representation of the training and transfer patterns for the transformational and random conditions. (Notice that the high level distortions are shared by both types of conditions.)

prototypes of three categories. The prototype provided an additional contrast among the conditions since the prototype was equidistant from all the old patterns in the five conditions. We also included new randomly generated exemplars of the three categories. In addition, we included foils, which were patterns that were not members of the three categories. The midpoint patterns between two successive transformations, which we will call gappers, were also shown. Gappers were included to test performance on new items embedded along a transformational path. Figure 1 gives a schematic representation of the patterns included in the transfer test as well as the learning set. The rings represent the various levels of distortion of the patterns. Note that this representation is of only one of the three categories and that the three high level distortions are actually at arbitrary distances from each other on the outer ring.

The subjects made categorization and recognition judgments to this set of patterns in an immediate test. Subjects returned to the laboratory 1 week later for a delayed test that also included these patterns. A delayed test permitted a separate analysis of how decay might interact with the different training conditions.

Expected Results

The following predictions are motivated by the idea that transformational knowledge is acquired when subjects see a progression of change within a category. For this transformational knowledge to be used, the transformation must be understood by the observer. Therefore, the change must be systematic and cumulative. The knowledge of this change causes the organization of the category such that things along the path are more salient. This increased salience of the path of change results in enhanced categorization and recognition performance.

The first set of predictions concerns contrasts between the transformational conditions and the random conditions. In general, we predicted an advantage for the transformational groups, relative to the random groups, in learning and subsequent transfer. These predictions are made by exemplar and feature-based models of classification but not by pure prototype models. A second set of predictions involves the relative performance of the different transformational conditions. That is, if subjects are acquiring transformational knowledge and using this knowledge in a categorization task, then systematic training should result in superior classification and recognition performance compared with random presentation of the transformational items.

One of the more interesting predictions concerns the gappers on the transfer test. Although research on scripts is semantically based, there is direct relevance to the present study. Much like the false recognition of the parts of the semantic scripts (e.g., Bower, Black, & Turner, 1979), we hypothesized that the gappers, or the implied states of the transformations, would be recognized as old (i.e., falsely recognized) at levels comparable to the training patterns by participants in the transformational conditions. That prediction is in line with models that allow similarity of the old instances to play a role in subsequent categorization judgments. Moreover, if subjects have more explicit knowledge of the transformation in the most ordered or systematic transformational condition, then we expect that the degree of false recognition of the gappers would be elevated in this condition when compared with the other transformational conditions.

Method

Participants. The participants were 80 undergraduate students at Arizona State University enrolled in an introductory psychology course. Participation in this experiment was for course credit.

Stimuli. The stimuli that were used in this experiment were nine-sided forms, similar to the dot patterns used by Posner, Goldsmith, and Welton (1967). These patterns appeared as white forms on a black background. The dots were connected in an arbitrary order that was kept consistent for all patterns in a category.

The learning set for all the conditions consisted of three categories of 15 patterns. For the transformational conditions these patterns were composed of three sets of five successive distortions of each of the three prototypes. The fifth distortion of any of these sets was a high level distortion. For these stimuli, a high level distortion is about 5.0 units/dot from the prototype and is approximately midway between a prototype and a random pattern (Homa, 1984).

For the two random conditions, each pattern in the learning set was distorted to a level equal to the level of distortion of a corresponding pattern in the transformational group. However, the distortions of each point of these patterns were not constrained to be in a certain direction and therefore the distortions bore no systematic relationship to one another. The three high level distortions were the same patterns used in the transformational conditions. Figure 2 illustrates the two types of learning patterns, as well as the prototype, for the two types of conditions.

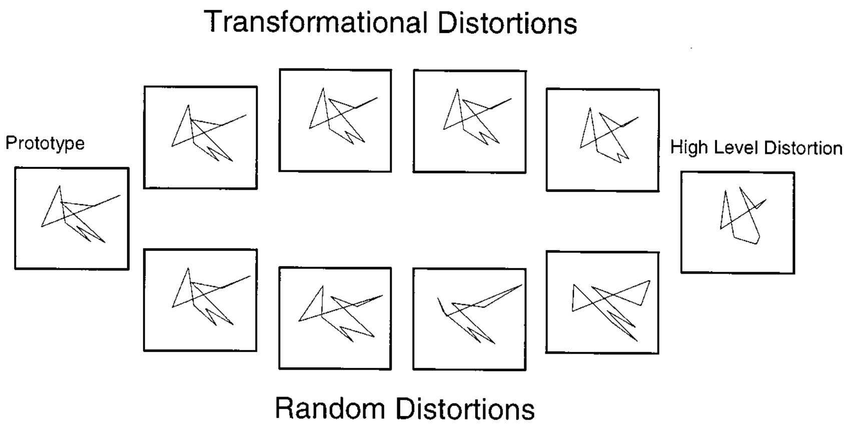

FIG. 2. An example of the patterns used in training in the two types of conditions. (The prototype is included in this diagram although this pattern was not shown to the participants during training.)

In the transfer set, a total of 111 old and new patterns were shown. In addition to the old stimuli, participants saw the three prototypes, 27 ''gappers'' (midpoint of two successive transformations), 27 new exemplars (low, medium, and high level distortions of the three prototypes), and nine foils (new random patterns).

Stimulus construction. To construct the transformational patterns and the gappers, the distance between each dot in the prototype and the corresponding dot in the high level distortion was calculated. For each dot, this distance was divided by ten and the result was added to the corresponding dot in each successive transformation. There were therefore nine patterns constructed between the prototype and the high level distortion. Four of these patterns were used as transformational patterns. The three patterns interspersed between each successive transformation were designated gappers. Two patterns remained that were not used in the present experiment.

The yoked patterns were constructed so that they maintained the same distances from the prototype as did the transformational patterns but the dots were free to move to any location within the allowed distance. The dots in the low, medium, and high level new exemplars moved, on average, 1.2, 2.8, and 5.0 units, respectively.

Procedure. Participants were randomly assigned to one of the five conditions. In each of these conditions, participants were given a series of study/test blocks that constituted the learning phase. Each of the study blocks consisted of serial presentation of the patterns in an order specific to the condition. Patterns were shown for 3 s each. The three groups of patterns (called A, B, and C) were block presented in each of the conditions, with patterns in group A appearing in a block of 15 followed by groups B and C. The subjects were told that all patterns within a category were different manifestations of the same object. The patterns were identified to the participant by the appearance of the appropriate letter in the upper right corner of the screen. The order of presentation of the categories was counterbalanced across the learning blocks. After all patterns for each of the three groups were shown, the participants were given their first test. Participants were shown all patterns in random order and were asked to indicate the category to which they believed each pattern belonged. Corrective feedback was given on every trial. This procedure was repeated for each of the study/test blocks.

The five conditions were the transformational systematic, transformational scrambled within radius, transformational scrambled within category, random with three blocks, and random to criterion. Participants in each of the first four conditions were given three study/test blocks.

(1) Transformational SYSTematic (T-SYST): The patterns in the learning trials for this condition consisted of forms on the three radii for each of the three categories. These patterns were presented sequentially within each series of five transformations, from low to high distortion.

(2) Transformational SCrambled within RADius (T-SC-RAD): The patterns used in the learning phase were identical to those used in the first condition. The patterns were block presented within each radius of five transformations. Within these blocks, the patterns appeared in random order. In effect, the participants saw five transformations (or states) on a radius but not in systematic order.

(3) Transformational SCrambled within CATegory (T-SC-CAT): The patterns used in the learning phase were the same patterns used in the other two transformational conditions. These patterns appeared in random order within each blocked category. For any particular participant, the same random order within a category was used on every study block.

(4) Random with an equal number of blocks (R-3BLOCK): The patterns used in the learning phase of this condition were yoked to be at an equal level of distortion to the learning patterns in the three transformational conditions. They were shown in a random order within each category. For a given participant, the same random order within a category was used on every study block.

(5) Random with criterion learning (R-CRIT): Participants in this condition saw the same patterns in the learning phase that were shown in the other random condition. We attempted to bring these participants to the same terminal level of learning as that of participants in the transformational conditions using learning curves obtained from a pilot study. These patterns appeared in random order within blocks of categories. For a particular participant, the same random order within a category was used on every study block.

After completion of the study/test blocks, participants were given instructions for the transfer test. Participants were asked to give two responses to each pattern. The first judgment was an old/new judgment with a confidence rating. The confidence rating scale ranged between one and three, with three indicating the most confidence. The second judgment was a classification judgment. The participants were informed that some patterns shown in the transfer test belonged to none of the categories and that the appropriate classification in such a case would be ''none.'' No feedback was given in this part of the experiment. One week after initial testing, each participant was given the transfer test once again. The procedure was identical to the one used in the immediate transfer test.

Results

Results are organized by learning, classification, and recognition performance, for contrasts between the transformational and random conditions, and among the transformational conditions. For each of the dependent variables, global tests for the five conditions are presented first, followed by contrasts among the transformational conditions.

Learning phase data. The mean probabilities of correct classification on the terminal learning trial were .963, .964, .950, .892, and .921 for the T-SYST, T-SC-RAD, T-RAD, T-SC-CAT, R-3BLOCKS, and R-CRIT conditions, respectively, $F(4, 75) = 4.35$, $MSE = .0035$, $p < .05$. A Newman–Keuls posthoc test indicated that the random condition with three learning blocks (R-3BLOCKS) was significantly different at the .05 level than the transformational groups, whereas the random group taken to criterion (R-CRIT), which required an average of 4.5 learning blocks, did not differ significantly from the transformational conditions. Therefore, the attempt to equate the degree of learning in the R-CRIT condition to the transformational conditions was at least partly successful.

Classification data for random and transformational conditions. Appendix A contains the mean observed probabilities of correct classification of all transfer items in each of the five conditions. An analysis was conducted for the overall accuracy of classification of patterns in both transfer tests across the two types of conditions (transformational and random). A significant main effect of condition was found, with the transformational groups showing an advantage on all patterns but the new instances, $F(1, 78) = 14.22$, $MSE = .065$, $p < .001$. The mean advantage (average across transfer test) for the transformational versus random conditions on the prototypes, old patterns, new instances, and foils was .014, .080, $-.103$, and .215, respectively. This effect is qualified by a three-way interaction between condition and type of pattern and delay, $F(3, 234) = 11.50$, $MSE = .018$, $p < .001$.

The mean probabilities of correct classification of the old patterns and the new exemplars in both transfer tests across the five conditions are displayed in Fig. 3. An analysis of the old patterns, as a function of two types of conditions, distortion level, and time of test revealed a significant advantage for the transformational conditions, $F(1, 78) = 12.79$, $MSE = .096$, $p < .001$, which increased as a function of increased distortion, $F(4, 78) = 12.02$, $MSE = .012$, $p < .01$.

In the analysis of the classification of the new exemplars, a significant main effect of condition was found, with the random conditions displaying an advantage, $F(1, 78) = 5.56$, $MSE = .220$, $p < .05$, which increased as a function of distortion, $F(2, 156) = 9.13$, $MSE = .033$, $p < .05$. Finally, an interaction of condition by time was found in which the transformational conditions showed improvement as a function of delay (overall .024 increase in accuracy), while the random groups showed a decline in accuracy levels (.023 overall decrease), $F(1, 78) = 6.22$, $MSE = .041$, $p < .05$.

The advantage of the random conditions in classification of new instances is probably illusory since the conditions differed dramatically in their assignment of foils into the categories. In effect, participants in the random conditions were far more likely to assign patterns, including foils, into the learned categories, especially on an immediate test. The mean rate of (erroneously) assigning foils into the learned categories was .649 for the random conditions versus .308 for the transformational conditions, immediate test; on the delayed test, these values were .472 and .384, respectively. An analysis based on signal detection theory suggested that in fact accuracy for the classification of new instances now slightly favored the transformational conditions ($d' = 1.91$ versus 1.88); for T-SYST, T-SC-RAD, T-SC-CAT, R-3BLOCKS, and R-CRIT these values were $d' = 2.12$, 1.86, 1.75, 1.82, and 1.94, respectively.[3]

[3] A similar analysis was done on the classification accuracy of the other types of patterns as well. However, there were no deviations from the patterns of results previously reported.

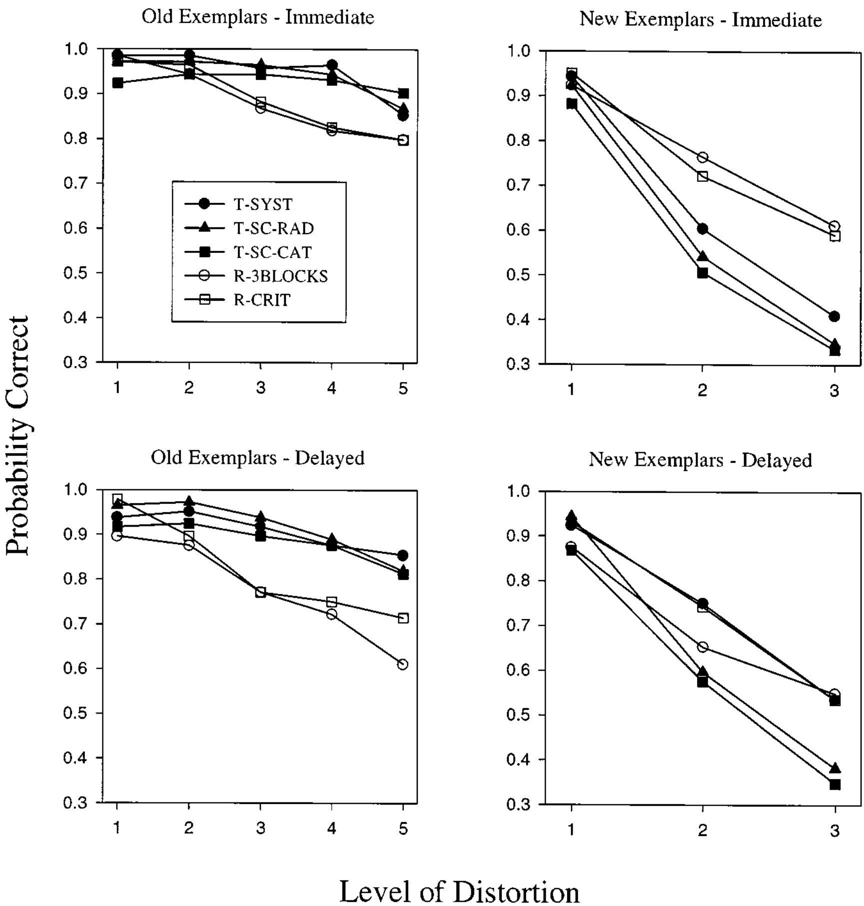

FIG. 3. The observed probabilities of correct classification of the old patterns and the new exemplars given by the five conditions in both transfer tests.

Comparison of classification among the transformational conditions. Comparison of T-SYST, T-SC-RAD, and T-SC-CAT revealed slight, but nonsignificant, advantages of the T-SYST and T-SC-RAD conditions relative to the T-SC-CAT condition in classification accuracy of the old, prototype, and gappers stimuli. As for the T-SYST and the T-SC-RAD conditions, although new patterns were consistently sorted more accurately following T-SYST training for both immediate and delayed tests, only the difference on the delayed tests proved significant; for medium distortions (T-SYST = .750; T-SC-RAD = .597), $t(30) = 1.83$; for high distortions (T-SYST = .535; T-SC-RAD = .382), $t(30) = 2.08$, both $ps < .05$, one-tailed test.

Recognition data for random and transformational conditions. The old/new ratings were combined with the confidence ratings to produce a six-point oldness scale. On this scale, a judgment of ''old'' combined with a confidence level of three yielded a value of six, ‘‘old’’ with a confidence of two was scored as five, and ‘‘old’’ with a confidence of one yielded a four. In contrast, a judgment of ‘‘new’’ combined with a confidence of one yielded a value of three, ‘‘new’’ with a confidence of two yielded a two, and finally ‘‘new’’ with a confidence of one was scored as one. The cell means of the oldness ratings of the various types of patterns given by the different conditions are contained in Appendix A.

An analysis of the oldness ratings as a function of condition, type of pattern (prototypes, old, new, foils), and the time of test revealed a significant interaction between condition and type of pattern, with the transformational groups providing the higher oldness ratings to all patterns except the new exemplars and foils, $F(12, 225) = 5.46$, $MSE = 6889.73$, $p < .001$. The transformational conditions had higher oldness ratings for the old patterns, $F(1, 78) = 12.41$, $MSE = 3.47$, $p < .001$. The decrease of oldness ratings as a function of time (collapsed over distortion level) was greater in the random conditions (0.48) than the transformation conditions (0.14), $F(1, 78) = 11.66$, $MSE = .446$, $p < .001$. Finally, the interaction between distortion level and condition was also statistically reliable; the random conditions showed a precipitous decline after the second distortion level, whereas the decline was gradual across the five levels of distortion for the transformational conditions, $F(4, 312) = 27.04$, $MSE = .270$, $p < .001$. Figure 4 shows the oldness ratings given to the old and new exemplars for the five conditions, shown separately for each transfer test.

Finally, the analysis of the oldness ratings of the foil patterns indicated a significant main effect of condition, with the transformational conditions being less likely to falsely recognize these patterns, $F(1, 78) = 17.97$, $MSE = .558$, $p < .001$.

Comparison of recognition among the transformational conditions. Figure 5 shows, for each transformational condition, the amount of forgetting as measured by a decrease in oldness ratings across the week delay for the old patterns as a function of the level of the training pattern. The values in Fig. 5 are differences between immediate and delayed oldness ratings. Overall, the amount of forgetting was least for the T-SYST condition (−.035), most for the T-SC-CAT condition (−.322), and intermediate for the T-SC-RAD condition (−.081), $F(2, 45) = 3.60$, $MSE = 0.53$, $p < .05$.

Analysis of performance on the gappers revealed a significant main effect of distortion level, with patterns furthest from the prototype being rated as less familiar, $F(2, 45) = 4.67$, $MSE = 17.35$, $p < .01$. However, contrary to our predictions, there was no main effect of condition, $F(2, 45) = 1.76$, $MSE = 181.38$, $p > .10$.

Discussion

Comparisons between the transformational and random conditions. Exposure to transformational knowledge enhances the learning of concepts. The transformational conditions exhibited a higher rate of learning compared to

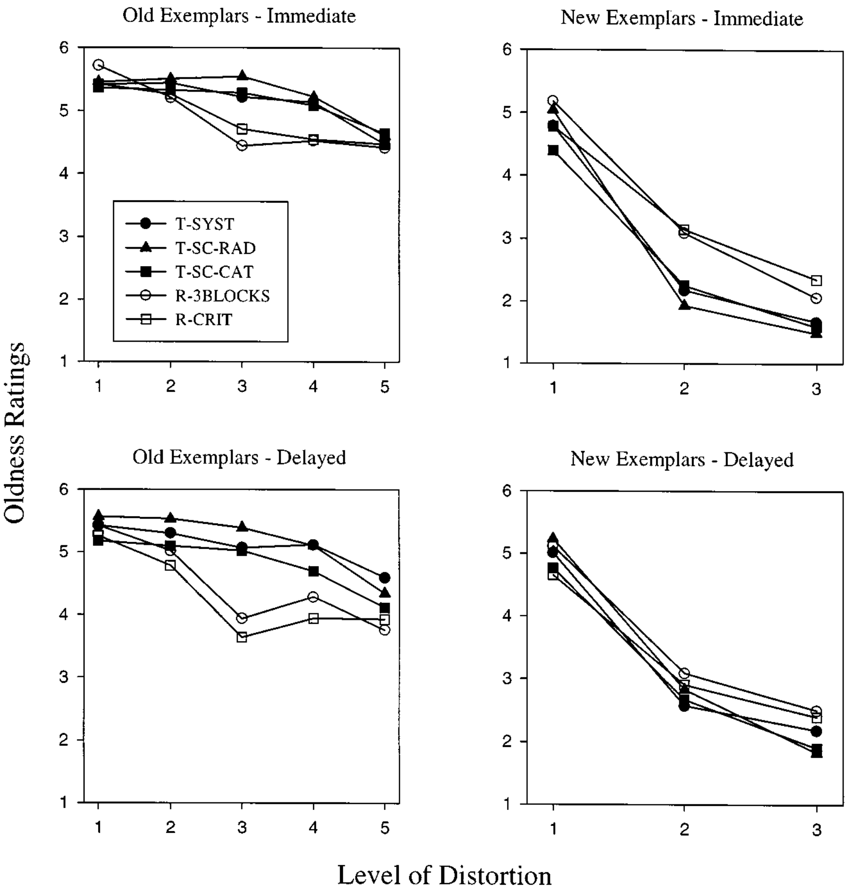

FIG. 4. Oldness ratings given by the different conditions to the old patterns and the new exemplars in both transfer tests.

the random conditions. Not only was the rate of learning superior for these participants, but the level of learning that they achieved could not be emulated by some participants in the random conditions with as many as eight blocks. These asymptotic levels of learning of the transformational patterns with so few exposures to the stimuli are scarce in the literature in experiments with nine-sided forms. (See, for example, Homa et al., 1991.) Although not the major focus of the present study, we believe the rapid and elevated learning in the transformational conditions is notable.

Similarly, the transformational conditions showed an overall advantage in classification and recognition performance. The transformational groups performed at higher levels of accuracy and were more confident in their recognition of old patterns in both immediate and delayed tests. The classification advantage increased at higher levels of distortion and also as a func-

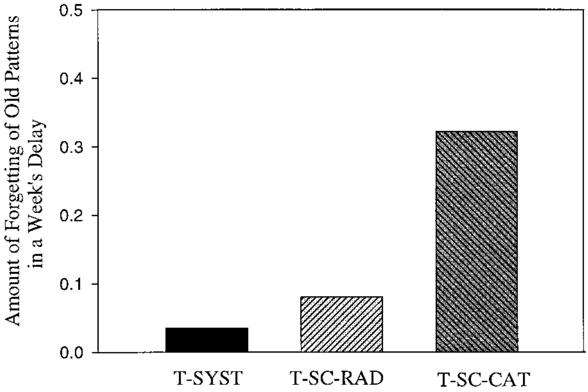

FIG. 5. The amount of forgetting of the old patterns across the week delay by each of the transformational conditions.

tion of delay. Participants in the transformational conditions were also more accurate in classifying the foils and were less susceptible to false recognition of these patterns. An apparent exception was the classification responses to the new instances, where the random conditions exhibited marginally superior performance on these patterns. However, closer inspection using a signal detection analysis suggested that the uncorrected performance was due to the adoption of a more liberal criterion, by participants in the random conditions, for incorporating patterns into the category. Evidently, participants who were exposed to transformations tended to be more conservative in terms of allowing new patterns into the categories. In sum, transformational training produced concepts that were more tightly knit and resistant to decay.

Comparisons among the transformational conditions. Although the effects of systematic versus random ordered training were generally in the right direction, only a few of the comparisons made among the transformational conditions reached statistical significance. For example, participants who were trained on the transformational items in a systematic order were more accurate at classifying new patterns than participants who were trained on the same patterns in a random order. Possibly, other effects were masked by levels of performance that were close to or at ceiling. Nonetheless, the trends, although not significant, were suggestive.

Our prediction that the gapper stimuli would be falsely recognized as old was confirmed, a result somewhat analogous to the false recognition of implied actions in scripts (Bower et al., 1979). In fact, these new stimuli received oldness ratings that were comparable to the training patterns, an outcome that did not change with a week's delay. Contrary to our predictions, false recognition of these stimuli did not vary among the transformational conditions. Specifically, we had anticipated that the systematic training condition would produce a memorial representation that was tightly sequenced, thereby precluding accurate oldness ratings of the gappers. Presumably, these sequenced events would be difficult to reassemble in the scrambled conditions, resulting in somewhat lower oldness ratings. Either our hypothesis was wrong or the effect of old–new similarity (comparable among the three transformational conditions) was sufficiently strong to mask potential differences in transformational knowledge. An alternative test, which might separate the influences of old–new similarity from transformational knowledge, would be to extend the transfer test beyond the final training stimuli, with the critical stimuli either falling on or off the training radius but at a comparable distance to the old instances. This was the purpose of Experiment 2.

EXPERIMENT 2

Experiment 2 attempted to separate two influences that may have contributed to the general advantage of the transformational conditions to the random conditions, as well as to the slight advantage of the Transformational Systematic relative to the Transformational Scrambled within Category condition. According to one view, the participant learns the transformational sequence and this knowledge enhances conceptual decisions. According to an alternative explanation, the participant simply stores the individual stimuli of the transformation but the transformation itself plays no role in subsequent category judgments. Rather, within-category pattern similarity is the critical component, i.e., at the time of test, the participant's judgment is based on the collection of stored instances (e.g., Nosofsky, 1988; Shin & Nosofsky, 1992). This view suggests that the average similarity between the test probe and the learning patterns is greater when the learning patterns fall along a transformational path than when they do not. This alone would be sufficient to explain the advantage of the transformation conditions relative to the random conditions. Although the advantage of the T-SYST condition relative to the T-SC-CAT condition for selected contrasts is inconsistent with this view (since the same patterns are used), the relative comparability of performance of the T-SYST and T-SC-RAD conditions is in line with this notion.

In Experiment 2, a direct test of these two potential influences for the Transformational Systematic condition was provided by pitting within-category similarity against transformational knowledge. The participants were trained on sequences of transformations in the first phase of this experiment. The participant's task in the transfer phase was not a standard categorization task but was instead to select that pattern which was most consistent with the previous sequence of patterns. The transfer test consisted of an array of three patterns, one of which continued the sequence (the transformational pattern) and two equidistant patterns that diverged from the sequence. The patterns in each test array were constructed such that each was equidistant from the nearest pattern of the learning set. A graphical representation of

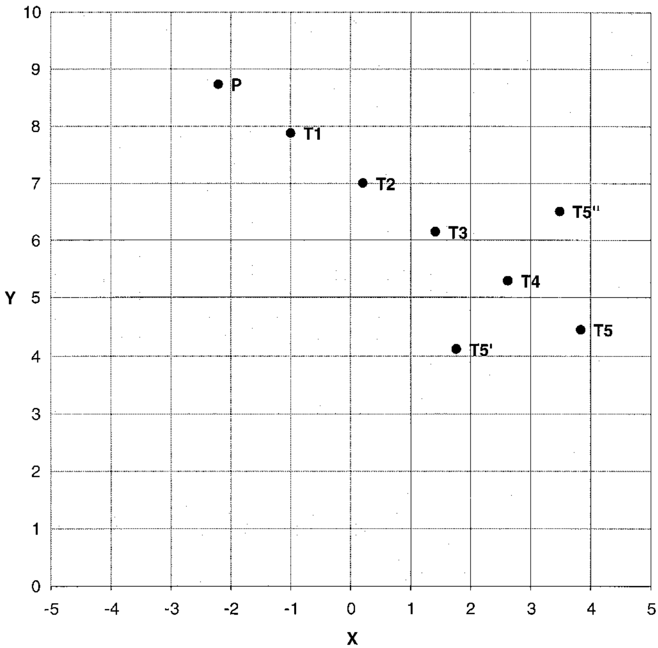

FIG. 6. Construction of a transformational path (T1–T4) from the prototype (P) to a high level distortion (T5), along with patterns displaced from the path but equidistant from the terminal T4 pattern (T5′, T5″). The *x* and *y* axes indicate only one of the nine (*x*,*y*) coordinates of a pattern, but the procedure was performed on all nine points of each pattern.

how these controls were invoked is shown in Fig. 6. Note, the two patterns that diverged from the transformational path were closer than the transformational pattern to the remaining patterns of the sequence. In short, if participants based their choice on pattern similarity, the test was biased against selection of the transformational pattern. That is, if participants preferred the transformational pattern, their choice must reflect continuation of the sequence and not similarity alone.

In Fig. 6, only one of the nine points for the learning sequence (prototype, T1, T2, T3, and T4) and the test array (T5, T5′, T5″) is shown; in the experiment proper, the procedure described here was performed for each of the nine points of a pattern. T5 is a high level distortion of the prototype that falls along the transformational path had the path been extended; T5′ and T5″ are the corresponding points for the other two patterns on the transfer test. Each pattern was constructed such that, for each of the pattern points, T5, T5′, and T5″ were equally close to T4 (the nearest pattern of the learning sequence for all patterns in the test array). In fact, the two patterns that diverged from the transformation fell on a line orthogonal to the transformational path, intersecting the path at the T4 pattern. In addition, T5 was further

TABLE 1
Interpattern Distances for the Learning Set (P, T1, T2, T3, T4) and Test Array (T5, T5′, T5″), Experiment 2

	P	T1	T2	T3	T4	T5	T5′
T1	0.98						
T2	1.95	0.98					
T3	2.93	1.95	0.98				
T4	3.91	2.93	1.95	0.98			
T5	4.89	3.91	2.94	1.96	0.98		
T5′	4.03	3.09	2.18	1.38	0.98	1.39	
T5″	4.03	3.09	2.18	1.38	0.98	1.38	1.95

from the remaining patterns of the learning sequence (P, T1, T2, and T3) than either T5′ and T5″. Table 1 shows the within-category distances of a typical test array (T5, T5′, and T5″) to each of the learning patterns (P, T1, T2, T3, T4).

Construction of the test array, described more fully in Appendix B, was applied to all nine points of each pattern, and therefore, the overall distance of the transformational pattern was identical to the two alternatives for the most similar pattern and less similar than the two alternative patterns for the remaining patterns of a sequence. The hypothesis was straightforward: if participants prefer T5 to T5′ or T5″, then transformational knowledge, and not pairwise similarity, is responsible for the judgment; if T5 is not preferred, then similarity, rather than transformational knowledge, is guiding performance.[4]

Two other points should be noted. First, as shown in Table 1, the similarity among the three test patterns in an array was quite high, corresponding to a low level distortion of each other (e.g., Homa, 1978; Posner, Goldsmith, & Welton, 1967). Due to this high level of similarity, we anticipated that slight differences would likely be obtained and we opted to run a sizable number of participants. Second, the procedure adopted for generating the test array actually ensured more than the pattern differences noted in Table 1: a comparable matrix for each of the nine points of a pattern would reflect the pattern shown in Table 1. For example, Table 1 shows that the average distance moved per dot between the prototype and T5 was 4.89 units; it was 4.03 units between the prototype and T5′ or T5″. These means are derived by averaging the nine points of each pattern, e.g., the distance for each of the nine corresponding points between the prototype and T5 was 4.58, 2.51, 9.26, 4.15, 1.46, 3.88, 0.93, 10.28, and 6.97; the corresponding distances for T5′ to the prototype were 3.76, 2.08, 7.64, 3.42, 1.21, 3.19, 0.76, 8.46, and

[4] This hypothesis assumes that psychological similarity is monotonically related to objective distance.

5.74. Note that each corresponding value was greater for the prototype-to-T5 contrast than for the prototype-to-T5′ contrast. In effect, each point of T5 was further from the corresponding points of the prototype, T1, T2, and T3 than was T5′ or T5″. Therefore, if the participant encodes each pattern as either composed of nine features (one for each point) or as an entire pattern, or any feature-combination intermediate between these extremes, the similarity advantage of T5′ and T5″ relative to T5 is maintained.

Method

Participants. The participants were 74 Arizona State University undergraduates, drawn from the same pool as in Experiment 1. None of the participants in Experiment 2 had participated in Experiment 1.

Materials and apparatus. The same types of patterns used in Experiment 1 were used in Experiment 2. In the learning phase, all patterns were shown via a Kodak Carousel projector. In the test that followed learning, participants were shown 18 patterns in booklet form with each pattern appearing on a separate page. In the transfer phase, participants were shown, via Carousel projector, the patterns corresponding to a given radius (e.g., P, T1, T2, T3, T4) and then turned over a page in a prepared booklet. Each page of the booklet contained a test array of 3 patterns, including the transformational pattern (T5) and two foils (T5′ and T5″). Each booklet contained 18 pages of test arrays.

Procedure. A modified Transformational Systematic procedure was used in the learning phase of Experiment 2. Participants ran in groups of two to eight and observed the projected patterns on a screen about 8–10 ft away. Each participant observed, in sequence, the prototype, T1, T2, T3, and T4 patterns for a given radius. Patterns were shown for 2 s, and six radii were presented for each category in succession, i.e., 30 consecutive patterns for a given category (e.g., category A). The procedure was repeated for another category (e.g., category B) and then the third category (e.g., category C). The entire procedure was then repeated and followed by a brief test. On the test, 18 patterns, 6 from each category, were quasi-randomly selected and presented in booklet form, with each pattern appearing on a separate page. The participant's task was to write the appropriate category label (A, B, or C) in response to each pattern. No feedback was given.

The transfer task followed immediately. As before, the participants observed a sequence of patterns from a given radius (prototype, T1, T2, T3, T4) but they were followed by an instruction to turn to the first page of the booklet. Each test array contained three patterns, including T5, with its position (left, middle, right) determined randomly. The participant was instructed to ''first choose the pattern you believe is the next step in the sequence that you have just seen.'' In addition, the participant indicated a goodness rating on a 10-point scale (1 = does not fit the sequence well; 10 = fits the sequence well) for each of the three patterns. The procedure was repeated for all 18 sequences (radii), again blocked by category.

Results

The mean rate of correct classification on the test following the two study blocks was moderately high (.777), with 39 of the 74 participants making no more than two errors.

We evaluated the likelihood of selection of the transformational T5 pattern in a variety of ways. First, the mean goodness ratings for T5, T5′, and T5″ patterns were 6.58, 6.18, and 6.06, respectively. The goodness ratings were significantly higher for the T5 patterns than for the average of the T5 and

T5″ patterns. $t(73) = 4.60$, $p < .001$. Second, 36 participants had mean goodness ratings that favored T5, 13 participants favored T5′, and 18 participants favored T5″; of the 13 participants with tied ratings, 5 favored T5, 5 favored T5′, and 3 favored T5″, z(approximation to binomial) $= 3.54$, $p < .01$. Third, an item analysis, computed on the 18 transfer stimuli (and collapsed across participants), again favored T5 relative to T5′ and T5″, $t(17) = 2.26$, $p < .05$.

Discussion

Experiment 2 demonstrated that when transformational knowledge and interpattern similarity are put into conflict, participants *can* select the transformational pattern. In fact, the impact of transformational knowledge might have been greater had pattern similarity been equivalent, rather than so favorable, to the alternative patterns. At least a few participants apparently relied on interpattern similarity since they performed considerably below chance in selecting the transformational pattern. Thus, one participant who never selected the transformational pattern (in 18 trials) was clearly below chance (6 of 18) and probably relied on interpattern similarity rather than transformational knowledge. Nonetheless, the results of Experiment 2 provide support for the hypothesis that transformational knowledge can be separated from interpattern similarity and that participants can select the transformational pattern at rates significantly greater than chance.

EXPERIMENT 3

Although Experiment 2 suggests that participants are able to access transformational information, we cannot assert that this information is normally utilized in a categorization task. An alternative hypothesis is that participants can access transformational knowledge when the instructions explicitly suggest it but that this knowledge plays no role in categorization judgments. Another concern is that the effects of order of presentation were not investigated in Experiment 2. If participants can use the knowledge of a transformation in a categorization task, then it follows that this transformational effect would be greater in the situation in which the stimuli that constitute a transformation are presented in a systematic order.

Experiment 3 explored the effects of systematic exposure to transformational stimuli akin to Experiment 1, but with modified transfer stimuli similar to those employed in Experiment 2. In Experiment 3, participants were trained on the transformations of three categories. All participants trained on the same items, but their training consisted of either systematic presentation within a transformational path or random presentation of the items within a category. After participants reached learning criterion, they received a transfer task in which they made speeded categorization judgments to patterns similar to those used in the test array in the second experiment.

Three additional modifications were also incorporated into Experiment 3 in an attempt to increase the strength of the transformational knowledge. First, participants were more thoroughly trained on the patterns so as to increase their knowledge of the transformation. In the previous experiments, there were multiple radii in each category with no indication of where one began and another ended. Therefore, the second modification was to train participants on only one radius per category with more transformational steps within the radius in order to make the transformation more salient. Finally, this radius extended further than the radii in previous experiments in hope that with such high distortion levels, the participants would rely more on the knowledge of the path to guide their judgments.

In addition, the distance of the test item from the end of the transformational path was manipulated. This was done for two reasons. First, we were uncertain how transformational knowledge and similarity might interact with distance. For example, one could argue that transformational knowledge might increasingly dominate similarity as distance to the test patterns increased. Alternatively, the impact of transformational knowledge might diminish as distances increased, thereby fostering greater reliance on pairwise similarity. Second, by manipulating this distance it was possible to decrease the similarity among the test items.

The speeded categorization judgments were used since the ceiling effects obtained in Experiment 1 may have masked performance differences among the transformational conditions. Although in assessing models of categorization, researchers have traditionally investigated accuracy of response, there has recently been more attention to the reaction time predictions that certain models might make (e.g., Ashby, Boynton, & Lee, 1994; Ashby & Maddox, 1994a; Lamberts, 1998; Nosofsky & Palmeri, 1997). However, even now, relatively few specific hypotheses can be found in the category literature to predict reaction times of categorization judgments. A noteworthy exception is the RT–distance hypothesis, which predicts that reaction time decreases as a function of distance of the stimulus in the perceptual space from the decision bound (Ashby & Maddox, 1994a). As the organization of the psychological space, including the location of the decision bounds, is unknown, it is difficult to test the predictions of this model in the present experiment. Another recent advance has been Nosofsky and Palmeri's (1997) Exemplar-Based Random Walk model of speeded classification. This model assumes that exemplars enter into a race to be retrieved in memory. These retrieval rates are a function of the similarity of the test items to items in memory. This model does not predict an advantage of the transformational systematic relative to the transformational scrambled order conditions, assuming constant similarity representations.

Instead, the prediction concerning response time was motivated by the literature that shows that participants react faster to stimuli that are expected than ones that are not (e.g., Rosch, 1975a; 1975b). For example, research in attention has demonstrated that stimuli are reacted to more quickly when participants are cued to the location of the upcoming target (Jonides, 1981; McLean & Shulman, 1978; Posner & Snyder, 1975). If, in fact, the participants have knowledge of the implicit path, then items located on this path should have higher expectancy values and participants should therefore be slower when categorizing the foils versus the equidistant transformational patterns. In addition, participants should be more accurate in classifying the patterns that continue the path.

The main prediction again concerns the variable of systematic training. If systematic training is fundamental to this process, then the advantage shown to the patterns that lie on the transformational path should be present only for the systematic condition and not for the scrambled condition.

Method

Participants. A total of 96 students at Arizona State University participated in this experiment as part of an introductory psychology course requirement. Each participant was randomly assigned to one of six conditions. These participants had not participated in the previous experiments.

Stimuli. There was one radius in each of three categories. In the learning phase the participants were trained on nine transformational patterns in each category (T1–T9). The transfer stimuli consisted of the three critical patterns for each category at each of three distortion levels of test items. There were 27 transfer items in total.

Stimulus construction. New foils were constructed using the principles of coordinate transformation. Basically, the idea was to place a hypothetical circle around each point in the T9 pattern. This circle had a radius of the average displacement per transformation. As shown in Fig. 7, the new foils (45°) were patterns whose points fell on this circle and were located at a 45° angle in either direction from the transformational path. This procedure created patterns that are off the transformational path, equidistant from T9, but were less biased in terms of similarity to the patterns earlier on in the path.

For each of three categories, stimuli were constructed so as to fall on only one radius or transformational path. These transformational paths, however, extended further than the paths

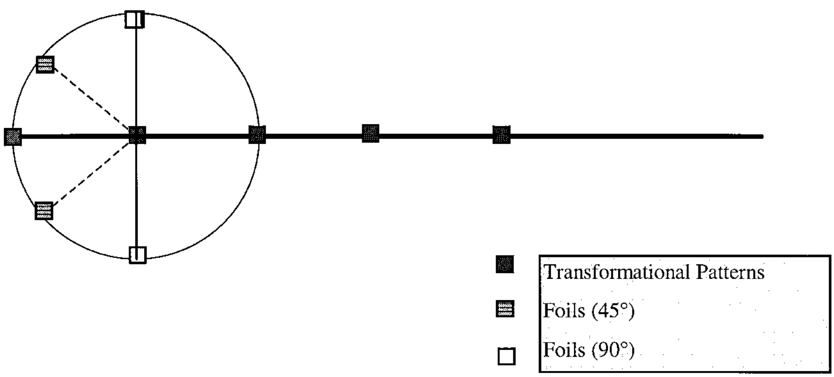

FIG. 7. A schematic diagram of the relationship between the 90° foils, the 45° foils, and the transformational patterns.

FIG. 8. A schematic diagram of one of the categories in Experiment 3.

in the previous experiments. More specifically, the learning stimuli in Experiment 3 spanned the distance equal to approximately twice a high level distortion (10 units/dot). In addition to patterns being further apart on the path they were also more numerous. There were nine patterns in each radius seen in the training phase as well as the set of test items seen in transfer.

The transfer test consisted of three sets of test items at various distances from the last item seen in training. Each of these sets contained two tangent patterns and a transformational pattern. The three patterns were equidistant from the last item seen in training. (See Fig. 8 for a schematic diagram of one of the categories.)

Procedure. The experiment was run on IBM-compatible computers using a program with 1-ms timing resolution. Each participant experienced five learning blocks, each of which included a study and a test phase, just as in Experiment 1. In the study phase, subjects were asked to simply observe the patterns that appeared with their appropriate category labels. Each pattern remained on the screen for 3 s. Each study phase was followed by a test phase in which participants categorized the training items and received feedback on each of their responses. The order of items in this phase was randomized.

There were six between-subject conditions which corresponded to the order of learning (systematic vs. scrambled) factorially combined with the distance of the test items from the last item seen in training (three levels). In the systematic conditions, participants trained on the patterns blocked within radii with the prototype first, followed by the next eight transformational patterns, in consecutive order.[5] In the scrambled conditions, the participant trained on the same patterns in a random order within each category. The random ordering of items within a category was kept consistent in each study block for a particular participant.

The learning phase was immediately followed by a transfer test in which participants were asked to categorize the previously unseen transfer items. The participants were tested at one of the three levels of the critical items. Participants were asked to make categorization judgments as quickly as possible, without sacrificing accuracy, by pressing a key indicating the category (A, B, or C) of each of the 27 transfer items. Subjects received no feedback in this part of the experiment.

Results

Participants in both conditions reached a high and equivalent level of learning by the fifth learning block (scrambled conditions = .968 and systematic conditions = .974). The transfer data were analyzed in terms of reaction

[5] Note that although the pattern referred to as the prototype was the prototype in Experiment 2, a recalculated prototype would be very distant from this pattern. It would lie in the middle of the transformational path.

time and accuracy performance. In this and all subsequent reaction time analyses, incorrect responses were discarded and reaction times more than twice the value of the mean of that condition were replaced with the cutoff value. In addition, reaction times under 100 ms were discarded. Finally, because of the lack of practice trials, the response to the first pattern that appeared in transfer was also discarded.[6,7]

The mean reaction time data are shown in the top panel of Fig. 9. The reaction time analysis did not yield a significant main effect of pattern type, $F(1, 120) = 1.36$, $MSE = 47624$, or organization of learning, $F(1, 120) = 2.15$, $MSE = 151210$, both $ps > .10$. However, there was a main effect of distance of test items from the training path, with mean reaction time increasing as a function of increased distance of the test items from the last pattern seen in training, $F(2, 120) = 6.10$, $MSE = 151210$, $p < .01$. This variable did not interact with organization of learning or type of pattern.

The interaction between pattern type and organization of learning was marginally significant, $F(1, 120) = 2.85$, $MSE = 47624$, $p < .10$. Subjects in the systematic condition responded 78 ms faster to the transformational than the tangential patterns, $F(1, 60) = 4.02$, $MSE = 48184$, $p < .05$. This contrast was not significant for subjects in the scrambled conditions, with mean reaction times to the transformational patterns slowed by 14 ms relative to the tangent patterns, $F(1, 60) = .14$, $MSE = 47065$, $p > .10$.

The accuracy data are shown in the lower panel of Fig. 9. There was a decrease of accuracy as a function of increased distance of items in the test array from the training patterns, $F(2, 120) = 3.02$, $MSE = .03$, $p < .05$. However, there were no main effects of pattern type, $F(1, 120) = .26$, $MSE = .007$, or organization of learning, $F(1, 120) = 1.41$, $MSE = .03$, both $ps > .10$. In addition, none of the interaction terms were significant (all Fs < 1).

Discussion

Experiment 3 demonstrated that given equal levels of accuracy, participants who were trained on the transformational stimuli in the sequence in which they transformed were faster at categorizing patterns that continued the path than patterns that diverged from the path. Participants who were trained on these same stimuli in a random order did not demonstrate this effect. In addition, there appeared to be an overall drop in performance with increased distance of the test items from the last training pattern. Interestingly, this variable did not appear to influence the magnitude of the transfor-

[6] When a response to a transformational pattern was dropped, the responses to the corresponding tangent patterns were also dropped in order to maintain the controlled distances from the transformational path. However, since there were two tangent patterns in every radius, when the response to one of the tangent patterns was discarded, the response to the transformational pattern was not. In this case the second tangent pattern maintained the necessary control.

[7] An analysis of the median reaction times yielded essentially the same pattern of results, with reaction times being on average 20 ms slower.

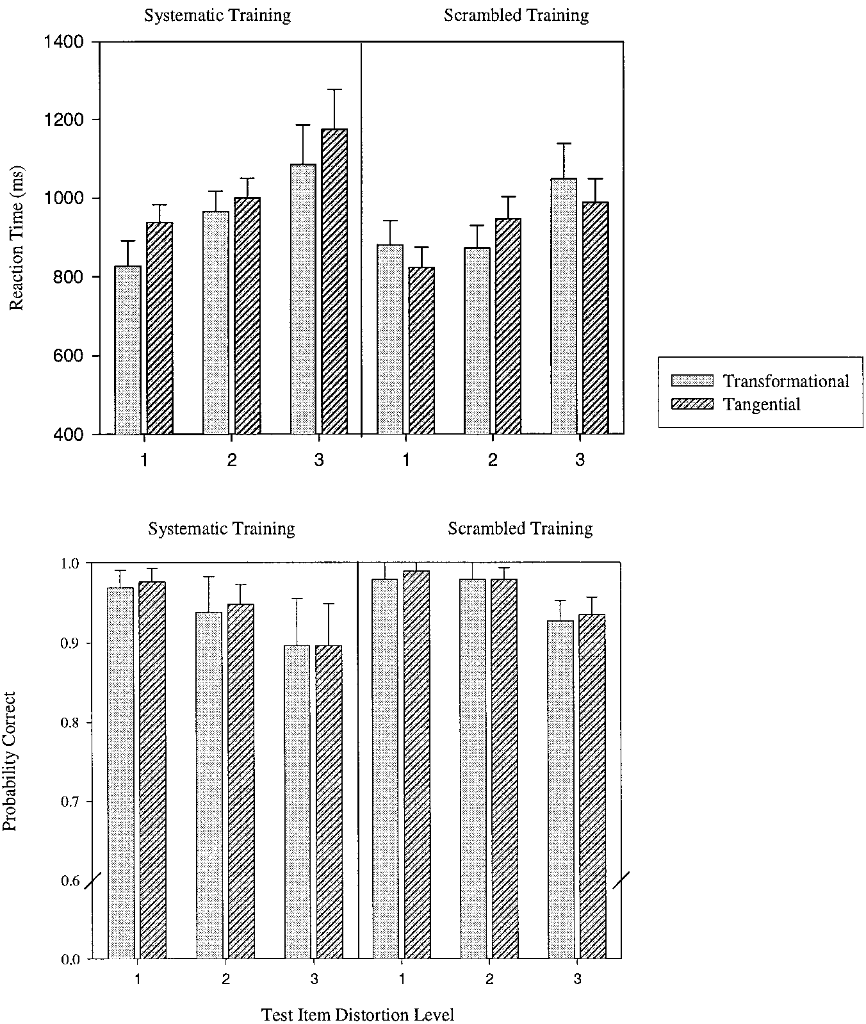

FIG. 9. Accuracy and mean reaction time performance in Experiment 3.

mational knowledge effect. The results of Experiment 3 provide support for the transformational knowledge hypothesis.

EXPERIMENT 4

Experiment 1 supported the idea that interpattern similarity can be separated from transformational knowledge. Experiment 2 demonstrated that participants could separate pairwise distance and transformational knowledge. Experiment 3 suggested that participants do use transformational knowledge in a categorization task if they are trained on these patterns in a systematic manner. Given the differences in performance between the scrambled and systematic conditions and assuming that interpattern similarity remains constant across the two learning situations, it is easy to argue that the various models of categorization cannot predict differences between these two conditions.

An interesting idea is that perhaps the interpattern perceptual similarity does not remain constant across both training procedures. The idea that concepts change as a function of experience is certainly not new (Goldstone, 1994; Hull, 1920; Nosofsky, 1986, 1987). Clearly, experience can modify our categories. Some theorists have argued convincingly that similarity and categorization are interdependent (Murphy & Medin, 1985). Therefore, as categorization performance changes, the underlying similarity relationships could change as well.

A number of studies point to the effects of making certain stimulus dimensions more salient. For example, Homa, Rhoads, and Chambliss (1979) investigated the effect of learning on similarity relationships. Participants gave pairwise similarity ratings for categorical patterns before learning, after learning to criterion, or after extreme category training. Multidimensional solutions were found for the similarity ratings and were compared across conditions. The results suggested that clustering of categories increases as a function of learning. Another example of increased salience of stimulus dimensions comes from studies of chick sexing. Distinguishing the sex of newly hatched chicks is an extremely difficult task, but with a high level of training sorters can reach levels of accuracy as high as 99.5% correct (Lunn, 1948, in Gibson, 1969) and can be trained to attend to the relevant stimulus dimensions (Biederman & Shiffrar, 1987).

The purpose of Experiment 4 was to use multidimensional scaling as a means of uncovering the psychological structure of categories formed after systematic, scrambled, and no training in the transformational categories. Multidimensional scaling (Kruskal, 1964; Shepard, 1962) is a technique which finds points in a multidimensional space so that distances in the space correspond to rated similarity. It is a tool that allows for modeling similarity in a way that uncovers relationships that are difficult to observe from raw similarity ratings.[8]

Participants either trained on the transformational stimuli in a systematic manner, the transformational stimuli randomly presented within a category,

[8] There are some limitations of MDS worth noting (e.g., Ashby, Maddox, & Lee, 1994b; Tversky, 1977). For example, Tversky argued that certain distance axioms, such as symmetry, are often violated in similarity ratings. Also, Ashby, Maddox, and Lee have demonstrated that averaging subject data can result in systematic changes in the computed MDS configuration. However, despite some of these problems, MDS still functions as a useful data reduction technique. Also, any bias should not favor one training technique over another.

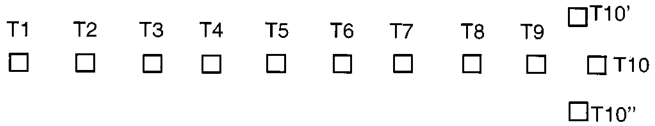

FIG. 10. A schematic diagram of one of the categories used in Experiment 4.

or received no training at all. In the rating phase that followed, participants gave pairwise similarity ratings to the training patterns in addition to some of the critical stimuli of Experiment 3. These critical stimuli were included so as to be able to determine whether the distance controls that were invoked in the physical space in Experiment 3 were maintained in a psychological space. Multidimensional scaling solutions were obtained for the three conditions.

One critical concern is whether the presumed transformational paths are realized in psychological space and whether they are better formed following systematic training. Specifically, there is no guarantee that the technique developed here for generating a sequence of patterns that systematically transform in *N*-dimensional space would be revealed as a clear path in psychological space. Furthermore, if the psychological spaces for the systematic and scrambled conditions are virtually identical, then the models have no a priori reason to suspect that similarity can account for results that suggest enhanced performance as a function of sequential learning. If, however, the similarity structure does change as a function of condition, then this can be taken as support for the idea that similarity and categorization are, to a certain degree, interdependent. It may also leave a way open by which models based on interpattern similarity can account for the results of Experiment 3 but would necessitate a more complex theory of similarity.

Another goal, as alluded to above, is to investigate whether the distance controls that were invoked in Experiment 3 are maintained in psychological space. An interesting possibility is that the critical stimuli will be rated as more similar to the last transformational pattern when they fall on the transformational path than when they fall off the transformational path. In other words, although the distances in physical space are controlled, perhaps the distances in psychological space are not.

Method

Participants. There were 30 participants who were predominantly graduate students at Arizona State University. Participants were paid $10 and randomly assigned to one of three conditions.

Stimuli. A subset of the stimuli used in Experiment 3 was used for this experiment. Figure 10 is a schematic diagram of the patterns used in Experiment 4. As in the training phase of Experiment 3, participants were shown nine patterns from each of three categories in the training or familiarization phase of this experiment. This training set was identical to the training set used in Experiment 3. The rating phase consisted of the training stimuli (T1–T9) in addition to the critical patterns (T10, T10′, and T10″) used in the third experiment for a total of 36 patterns. These critical stimuli consisted of the three test patterns (two 45° foils and one transformational pattern) for each category.

Procedure. For participants in the two learning conditions, the training procedure was identical to the procedure used in the third experiment. Participants were taken through five study/test blocks and, as in Experiment 3, participants either trained on the patterns in a systematic or scrambled fashion. In the no-learning condition, the training phase was replaced by a familiarization phase in which the participants were allowed to view all the training patterns once, for 3 s each, in a random order, but were given no information about any of the categories. This was done in order to familiarize these participants with the range of patterns so that they could gauge their rating scale accordingly.

After this initial phase was complete, the participants were asked to make similarity judgments to the stimuli. Each possible pair of stimuli appeared on the computer screen with left/right placement randomly determined and stayed on the screen until the participant made a response. In addition to the stimuli, a nine-point scale appeared on the screen to remind the participants of the range to use. Participants were instructed to use as much of the scale as possible to indicate the different shades of similarity, with ''9'' indicating extremely dissimilar stimuli and ''1'' indicating very similar stimuli. The trials were self-paced, and the whole session lasted approximately 1 h on average.

Results

Multidimensional scaling (MDS). For each of the learning conditions, the data matrices were averaged and the resultant matrix was used to find a scaling solution. A program called KYST (Kruskal, Young, & Seery, 1973) finds solutions in *N*-dimensions while minimizing a value called stress. Stress is essentially an indication of the fit of the model that is based on normalized residual sums of squares. Although a more dramatic drop in stress values occurred at two dimensions, the addition of a third dimension improved the fit considerably (the range of stress values [Kruskal's (1964) stress formula 1] for three dimensions was .054 to .08). In addition, a three-dimensional fit still allows a visual analysis of the data that higher dimensions do not. For these reasons, three-dimensional solutions were chosen for further analysis.

The resultant spaces provided by the three learning procedures can be compared in their overall organization. Generally, the spaces are very similar. A program called CONGRU (Olivier, computer program) was used to rotate the spaces so that they are maximally similar by minimizing sum of squared deviations. The program then computes a Spearman correlation coefficient indicating the degree of correspondence of the *x* and *y* values in the different solutions. The resultant ρ values from this analysis were .993 for the scrambled and systematic spaces, .980 for the systematic to no-learning mapping, and .985 for the no-learning to scrambled mapping. Figure 11 shows the three-dimensional plots for the three conditions, rotated to be maximally similar.

Computing structural ratios for the different learning conditions provides a measure of the degree of organization of these spaces (Homa et al., 1987). These ratios were obtained by calculating the distance from each pattern in the space to other patterns in the same category and dividing that distance

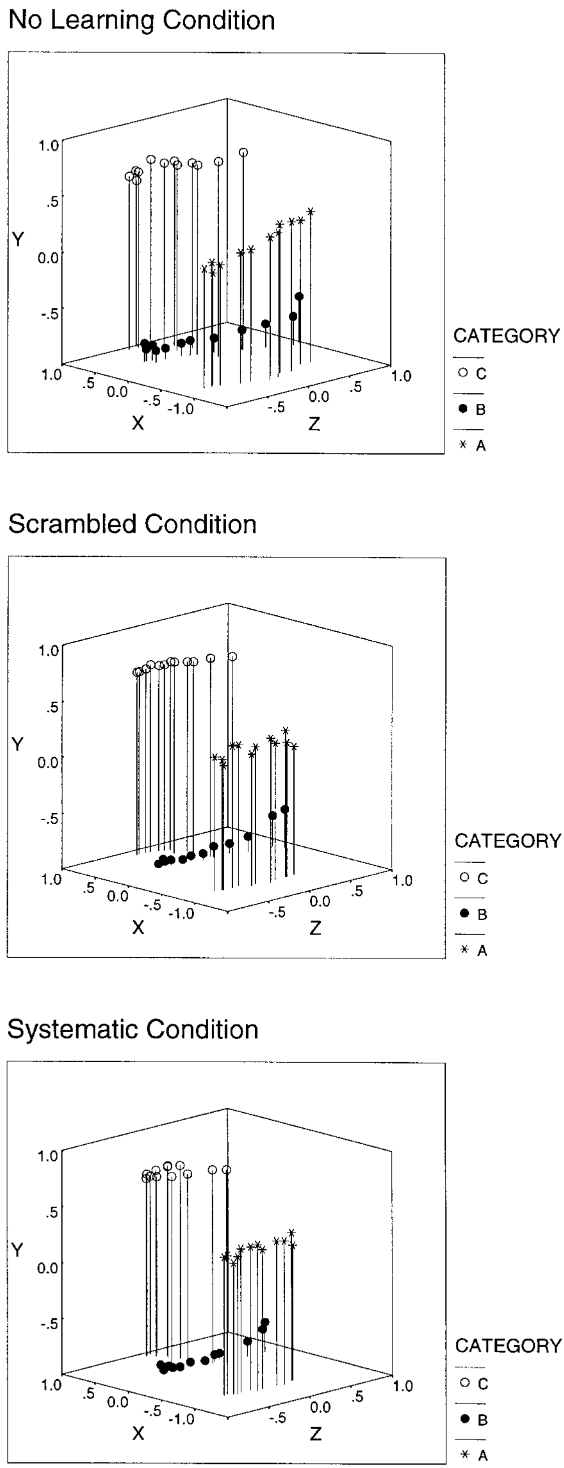

FIG. 11. MDS solutions for the three conditions.

by the distance to the patterns in the other categories. Smaller ratios indicate a greater degree of structure. This ratio was larger in the scrambled condition (.429 within/1.69 between = .254) than in the systematic condition (.402 within/1.69 between = .237), suggesting an increased organization of patterns in the systematic condition. Consistent with expectations, this ratio was highest for the no-learning condition (.571 within/1.658 between = .344). Since the distributional characteristics are not clear, we do not report any statistics for these analyses.

A more fine-grained analysis consisted of examining distances between the last transformational pattern seen in learning (T9) and the critical stimuli (T10, T10′, and T10″) for the three conditions. One possibility is that the distance between T9 and T10 will be smaller than each of the distances between T9 and the foils (T10′, T10″), despite what would be predicted by physical distances. In addition, this discrepancy between physical and psychological distances may be greater for the systematic condition than the scrambled condition. (Refer to Fig. 10 for a clarification of the relationship among these patterns.)

The top part of Table 2 gives the values of these comparisons for the three conditions. Physical similarity alone predicts that the distances in the two columns should be equal. However, contrary to what similarity-based models might predict given the ordering of mean reaction times in Experiment 3, except for the no-learning condition, the distance between T9 and T10 seems to be greater than the distance between T9 and the foils. Interestingly, this difference is greater for the scrambled condition than for the systematic condition. One must caution, however, that the reliability of these findings is

TABLE 2
Fine-Grained Comparisons in the Computed KYST Configurations

Condition	Foils (T10′, T10″)	Transformation (T10)
	Distance to last training item (T9)	
Systematic	.085	.107
Scrambled	.113	.163
No-learning	.100	.096
	Average distance to all within-category training items (T1–T9)	
Systematic	.459	.469
Scrambled	.497	.556
No-learning	.658	.678
	Average distance to all between-category training items (T1–T9)	
Systematic	3.13	3.11
Scrambled	3.13	3.16
No-learning	3.08	3.16

unknown because it is not possible to do a statistical test with only three measurements (one for each category) of each of the comparisons. An analysis of the raw similarity ratings, which does allow for statistical comparison, produced the same general pattern of results, with no significant differences between the similarity ratings of the two types of patterns and the training sequence for any of the conditions.

Another comparison included in Table 2 is the distance of the critical items (foils and transformational patterns) to all the other items in the path. These distances were calculated by averaging the distance of the items in the path to the two kinds of patterns. These values are an indication of the ‘‘perceptual length’’ of the path. Note that, consistent with the calculated structural ratios, the transformational path appears shortest for the systematic condition. As these are not an exact measure of the length of the path, the distance between the first and last training items was calculated to confirm this finding. As expected, this distance was shortest for the systematic (0.843) condition followed by the scrambled condition (0.893) and the no-learning condition (1.183).

INDSCAL analysis. Each of the individual participant matrices was next entered into a program called INDSCAL (Carroll & Chang, 1970). INDSCAL has one major advantage over other MDS programs. In addition to providing the *group stimulus space*, in which the stimuli are represented in terms of their similarity to one another, it also provides a *weight space*, in which each subject’s dimension weights used in the scaling are given. In a sense, a plot of these weights represents the similarity of each participant’s ratings with regard to other participants in terms of their shared dimensions of perceptual analysis. INDSCAL allowed us to investigate whether the participants in the different conditions were likely to cluster as a function of training procedure.

Each individual participant’s matrix for the three conditions was entered into the program. The resulting subject weight space in three dimensions is depicted in Fig. 12. Notice that the participants do seem to cluster as a function of training procedure. One way to quantify this apparent clustering is to first calculate the distances from each participant to other participants in the same condition as well as the distance for each participant to participants in the other conditions. Participants were closer within a condition (.146 units) than between conditions (.154 units).

Discussion

In general, the psychological spaces were very similar across conditions. The spaces were highly structured, with clearly visible transformational paths. There was even a surprising degree of structure in the no-learning condition. It appears that much of the transformational event can be reconstructed without explicit knowledge of the categories or exposure to the

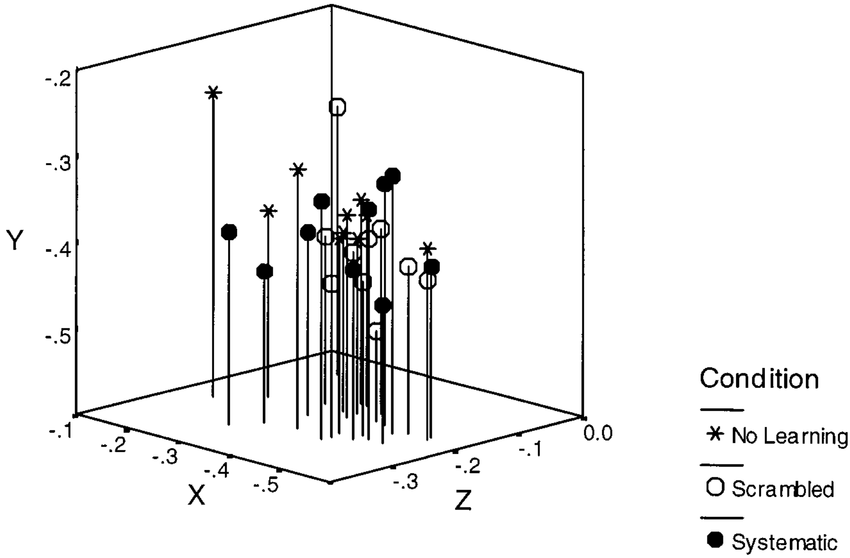

FIG. 12. The INDSCAL subject space.

transformation. That is, the perceptual similarity between items is enough to reconstruct the transformational order.

The results of Experiment 4 support the notion that similarity is modified as a function of training procedure. Despite the consistency across conditions, there were small but significant differences in the expected directions. The structural ratio analysis and the INDSCAL analysis indicate that participants in the systematic conditions have different similarity relationships than participants in the other conditions. Systematic training provided for more structured categories, with within-category distances being smaller than in the scrambled condition. In addition, systematic training resulted in the shortest path. The no-learning condition showed the highest structural ratios, indicating that as learning progressed, the categories become increasingly structured.

Finally, the fine-grained comparisons allowed an examination of how the critical patterns that were used in the last two experiments were perceived as a function of training procedure. These comparisons are important in that they provide the basis of predictions by models that rely on interpattern similarity. Two basic findings emerged. Although this trend was not significant in an analysis of the similarity ratings, the physical distances were more closely matched in the no-learning condition than in either of the learning conditions. Second, despite faster mean reaction times to the transformational patterns in the third experiment by subjects in the systematic training groups, these patterns maintained the controlled distances from the last item seen in training.

GENERAL DISCUSSION

Basic Findings

The goal of the present research was to investigate the importance of exposure to transformations in category learning. Objects in the environment evolve in a way that is principled. We hypothesized that exposure to this principled change should enhance categorical knowledge.

Experiment 1 demonstrated that learning is speeded, classification is enhanced, and forgetting is reduced for participants trained on transformational patterns versus those trained on random patterns equated for distance to the category prototype. Although many differences between the transformational conditions in Experiment 1 were in the right directions, the performance of these subjects was high and the differences were slight and only significant on a few comparisons. One advantage of systematic training was in the classification of new patterns. Participants who were trained on the transformational items in a systematic order were more accurate at classifying new patterns than participants who were trained on the same patterns in a random order. Another advantage was in the recognition of the old patterns with the systematic training providing higher levels of recognition of the training patterns.

Experiment 2 demonstrated that participants *can* access explicit knowledge of the transformation. After training on systematically distorted sequences, participants were able to pick out the next step in the sequence from an array of patterns in which this step was the least similar to the training instances. This indicated that participants were not guided solely by similarity judgments, but were able to use transformational knowledge when instructed to do so. This result also suggests that the cognitive representation of the transformation is a viable construct, separate from pairwise pattern similarity.

Experiment 3 demonstrated that participants, trained on principled change, categorize novel patterns more quickly if they lie on the transformational path. Importantly, this outcome was only obtained for subjects training on transformational patterns in a systematic order. When training occurred with the same distortions in random order, no advantage in speed of response occurred to path-continuing versus path-violating patterns. It is possible that systematic training creates expectancies of forthcoming patterns that are then more quickly classified.

Experiment 4 showed that the effects of systematic transformational training are readily apparent in a psychological space, as determined by multidimensional scaling. In particular, transformational paths are shorter and the entire categorical structure is more structured when participants train on patterns in a systematic manner. In addition, the INDSCAL analysis indicated that there were systematic differences among the conditions as manifested by clustering of participants in the weight space. Finally, fine-grained analyses revealed that the patterns that were controlled in terms of their distance from the path in a physical space maintained this control in the psychological space. The patterns that continued the transformational paths versus control patterns that diverged from the path were roughly equidistant to the most similar training pattern. Therefore, the speed advantage of the systematic condition on the transformational patterns in Experiment 3 was not due to a psychological shortening of the transformational paths.

A Theory of Transformational Knowledge

These results suggest that the human mind is receptive to and can profit from systematic changes in patterns. This ability is fundamental. It is necessary, for example, for identification of constantly changing objects; for recognizing an old friend after time has significantly changed her appearance; and it is essential for being able to see that tadpoles and frogs belong to the same category. The work of Seamon and his colleagues (Seamon, 1982; Seamon, Stoltz, Bass, & Chatinover, 1978) provides evidence that participants are able to recognize faces that have changed dynamically. They argue that it is a cognitive ability that occurs in both incidental and intentional learning. We contend that this ability is not limited to faces or any other ‘‘special’’ stimuli, but is about the knowledge of permissible transformations. We further argue that this knowledge is crucial to our understanding of objects in the environment.

This combination of what must be both extrapolation and interpolation processes causes heightened status for things along the path of change and, at least implicitly, for the path itself. As a result, objects that continue the path are expected, and when they are seen, categorization judgments proceed faster. The sensitization of the path also makes things that are on the path seem more similar, as was evidenced by the shorter paths in the systematic condition.

Others have argued for phenomena that resemble this process. The idea that people can extrapolate beyond their experiences is well documented (e.g., Brehmer, 1974; Koh & Meyer, 1991). For example, Koh and Meyer demonstrated that participants in their experiments could induce certain functions involving stimulus-response relations well beyond what they had experienced in learning. In addition, although the focus of this paper has not been to examine changes in memory as a function of implied momentum, our findings certainly have some parallels to the representational momentum literature in that both phenomena are tied to perceiving the course of change. However, there are also some fundamental differences. First, the representational momentum effect is found only in small temporal durations after the last item is seen. Second, if the test items are more than slightly distorted, the effect disappears. These two differences indicate that although these appear to be related phenomena, they may draw on somewhat different processes.

The question of how information about transformations is developed and represented is an interesting one. Any processing account must include an explanation of the importance of sequencing of the items to transformational knowledge. One possibility is that transformational knowledge is induced by the co-occurrences of similar items in memory early on in the transformation. This makes higher order invariances of the stimuli more salient thereby guiding expectations of upcoming stimuli. The notion that commonalties between stimuli seen early in training guide future categorization performance has recently received support (Ross, Perkins, & Tenpenny, 1990). In addition, Medin and Bettger (1994) have demonstrated that sequencing items so that there is minimal change between any two consecutive items results in better recognition of category items. Therefore, in this sense, the transformational knowledge is a function of the display of information in a sequenced way. Transformational knowledge is not only about the sequencing of stimuli, however, because subjects in the systematic conditions are only faster in classifying the items that continue the path rather than being faster on all items.

As for the representation of transformational knowledge, one possibility is that knowledge of transformations allows for the development of additional features or dimensions of objects along a transformational path, as suggested by perceptual learning theory (Gibson, 1969). For example, Goldstone and colleagues have recently demonstrated that participants in laboratory experiments can be sensitized to both existing (Goldstone, 1994) and arbitrary (Goldstone, Steyvers, & Larimer, 1996) dimensions.

Implications for Theories of Categorization

Findings of enhanced categorization performance as a function of exposure to principled change are important to understanding the categorization process in general and, therefore, to any theory of categorization. Not only has this variable gone unstudied, it indicates a necessity to consider a higher order of organization among the stimuli.

There have been some recent advances in developing models that make predictions of the timing of categorization judgments (e.g., Ashby & Maddox, 1994a; Lamberts, 1998; Nosofsky & Palmeri, 1997) but none of the timing models predict an advantage of the transformational patterns for the systematic condition. Instead of assessing specific models, we can assume that reaction time instantiations of the models follow the general flavor of the representation models on which they are based (i.e., that the exemplar models would require only knowledge of the similarity relationships between exemplars or that prototype models would predict response times as a function of distance from the prototype). This type of approach was recently taken by Ashby, Boynton, and Lee (1994).

Given these assumptions, the data indicating that only the systematically trained groups were faster in classifying the transformational patterns would not be predicted by any of the models. Take, for instance, exemplar-based models. At first glance, the data showing modified similarity as a function of learning seems to support this group of models. Perceived similarity is systematically different across conditions, as would be necessary for these models to predict the obtained reaction time results. However, the similarity ratings do not capture the entire story. The transformational patterns are in fact further apart (although not significantly further) from the last item in the training set than are the control patterns, even in the systematic condition. Although this was more evident in the scrambled condition than in the systematic condition, it still contradicts the ordering of the mean reaction times. A model that considers the order of presentation of stimuli, such as the Rational Model of Classification (Anderson, 1991), may be able to make different predictions for the conditions solely as a function of order. That is, an iterative algorithm that is order-sensitive might result in categorical structure that could vary based on the order of instance presentation. Indeed, order effects have been documented in the literature (e.g., Medin & Bettger, 1994). However, a model such as this would not be able to predict the advantage for the systematic conditions only on the transformational patterns. What seems to be necessary is a model that takes advantage of redundancies in close temporal slots. More research is needed before such a model can be formalized and tested.

Similarity and Categorization

The results and theoretical perspective of the present study are relevant to recent concerns regarding similarity and its sufficiency in explaining categorization or conceptual coherence (Medin, Goldstone, & Gentner, 1993; Murphy & Medin, 1985; Rips, 1989). The focus of Experiment 4 was to investigate the possibility of fundamental changes in the similarity relationships as a function of learning the transformational categories. Yet, researchers generally attempt to explain categorization behavior in terms of similarity relationships that are measured after categories are learned. Given the flexibility of similarity, its utility as an explanatory construct has been questioned. For example, Murphy and Medin make the argument that similarity and categorization are considerably interdependent and that similarity may be more useful as the dependent rather than the independent variable. It appears from Experiment 4 that the similarity relationships across the three conditions have a lot in common. However, there are significant systematic changes in the psychological space as a function of learning the categories, suggesting some interdependence between similarity and categorization.

Does this mean that similarity is a vacuous construct? We argue that although similarity and categorization may at times be interdependent, that is not to say that similarity cannot be useful. Clearly, even with changes in the learning procedure, there was a core of similarity relationships that was common across the three conditions. Although some would argue that there are no perfect taxonomies (e.g., Lakoff, 1987), others have demonstrated (e.g., Sneath & Sokal, 1973) that animal and plant classification can be functionally achieved via unambiguous physical measurement such as skeletal properties (bone length and angles). For example, Wood (1982) has provided a demonstration that the 17 species of stork can be recaptured by precise measurement of numerous characteristics (skull, humorous, sternum, etc.) from these species. Furthermore, ''classification based on skeletal morphology are highly congruent with those recalculated with behavioral data'' (Wood, 1982, p. 88). Thus, similarity can be a useful construct, but a clear operational definition of similarity is necessary. We advocate that the problem lies in providing an adequate, noncircular definition for similarity or noting that its implementation may take on different forms depending upon the particular stimulus domain (a view recently expressed by Heit & Rubinstein, 1994). We are sympathetic with Murphy and Medin's (1985) argument that ''. . . the notion of similarity must be extended to include theoretical knowledge'' (p. 291) and that ''. . . much of our reasoning about concepts may be based on constraints about operations that are permissible'' (p. 295). The present study is an attempt on our part to capture permissible operations that are not exclusively based on physical similarity among specific members. However, it is clear from findings of this and other studies (Homa, Rhoads, & Chambliss, 1979; Murphy & Medin, 1985; Rips, 1989; Rips & Collins, 1993) that using similarity to predict the categorization is explanatory only at a certain level. Perhaps more focus could be placed on investigating the types of variables that affect both similarity and categorization instead of simply explaining changes in categorization behavior as being a result of changes in similarity relationships. Hopefully, the present study, with its focus on categorization and similarity relationships as a function of transformational knowledge, represents a step in the right direction.

Representation

The potential performance differences among the transformational conditions may be less serious for theories of category representation than a theoretical dilemma raised by the transformational conditions themselves—what, precisely, is the representation of a category? To take exemplar models as an example, a tacit assumption of all studies exploring these models of classification is that each manipulated pattern results in an encoded exemplar, perhaps modified by bias or attentional demands. If that stimulus is shown a second or third time, different exemplar representations are likely formed. However, this perspective is mute to the very real issue of how one should define an exemplar which changes its appearance in real time. For example, what is the exemplar or exemplars when one observes a face that changes expression over some time period, such as 20 s? Logically a single exemplar representation cannot be formed in this period unless that exemplar is either a weighted average over the time period or the exemplar is the most salient representation during this interval. If the former, then the stored exemplar is really an abstraction; if the latter, then the argument must assert the dubious claim that the remaining potential exemplars are neither selected nor stored for representation. If the argument is made that multiple exemplars are formed during the time period (say 20 s), then it becomes necessary to posit a mechanism similar to the perceptual moment hypothesis, (e.g., every 100 ms or so, a different exemplar is encoded or stored). If we accept this hypothesis that multiple exemplars are stored during continuous viewing, then two problems emerge: what are the time units for formation? And are these exemplars also abstractions over a shorter interval? Regardless, models of classification make the tacit assumption that an exemplar is simply an internalized representation of the stimulus, without addressing the problems inherent with the reality of changing expressions and time. Change of the representation over time is a serious issue for all models of categorization but more so for models which do not incorporate abstraction processes.

Future Directions

An interesting idea is brought on by the contrast to the representational momentum research (e.g., Finke & Freyd, 1985; Freyd, 1987; Pinker, Choate, & Finke, 1984) and work that Kuhl (1991) has done regarding the prototypes of vowel categories. In her experiments, Kuhl asked participants to indicate when a referent sound had been changed. The participants heard sounds that were on a vector from the prototypical stimuli to a non-prototypical stimuli. The only difference between the two groups was the direction in which the stimuli were presented. Her hypothesis and results suggested that the prototype acts as a perceptual magnet so that surrounding members are perceptually assimilated into it to a larger degree than would be predicted by psychophysical distance alone. In support of her hypothesis, she found that to hear a difference, the participants in the prototype referent group had to go further away from the referent than in the non-prototype group.

Intrinsic to both these research paradigms is the directionality of the effect. The representational momentum effect, by definition, requires it to occur only in the direction of change, as does the perceptual magnet effect. However, in the case of transformational knowledge, if the subject is acquiring knowledge of the path of change, then perhaps this effect should manifest itself in either direction. Seamon (1982) indicated that participants display the ability to recognize faces that have changed in either direction. Perhaps this sort of bidirectional finding could increase our understanding of the processes involved in this phenomenon.

APPENDIX A
Classification and Recognition Cell Means for the Five Conditions in Experiment 1

	Classification				Recognition			
	Immediate		Delayed		Immediate		Delayed	
Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Old patterns: Distortion level 1 (O1)								
T-SYST	0.986	0.038	0.938	0.181	5.4	0.6	5.4	0.7
T-SC-RAD	0.972	0.111	0.965	0.067	5.5	0.6	5.6	0.5
T-SC-CAT	0.924	0.097	0.917	0.131	5.4	0.4	5.2	0.8
R-3BLOCKS	0.986	0.038	0.896	0.192	5.7	0.3	5.4	0.5
R-CRIT	0.972	0.064	0.979	0.045	5.4	0.6	5.3	0.5
Old patterns: Distortion level 2 (O2)								
T-SYST	1.000	0.000	0.951	0.140	5.4	0.8	5.3	0.9
T-SC-RAD	0.972	0.086	0.972	0.050	5.5	0.6	5.5	0.5
T-SC-CAT	0.944	0.091	0.924	0.120	5.3	0.6	5.1	0.9
R-3BLOCKS	0.944	0.115	0.875	0.202	5.2	0.7	5.0	0.8
R-CRIT	0.965	0.067	0.896	0.154	5.3	0.5	4.8	0.7
Old patterns: Distortion level 3 (O3)								
T-SYST	0.958	0.090	0.917	0.179	5.2	1.0	5.1	0.9
T-SC-RAD	0.965	0.088	0.938	0.099	5.5	0.5	5.4	0.6
T-SC-CAT	0.944	0.081	0.896	0.125	5.3	0.7	5.0	0.7
R-3BLOCKS	0.868	0.182	0.771	0.267	4.4	0.9	3.9	0.8
R-CRIT	0.882	0.137	0.771	0.255	4.7	0.9	3.6	1.1
Old patterns: Distortion level 4 (O4)								
T-SYST	0.965	0.067	0.875	0.210	5.1	0.9	5.1	0.7
T-SC-RAD	0.944	0.091	0.889	0.152	5.2	0.6	5.1	0.8
T-SC-CAT	0.931	0.098	0.875	0.128	5.1	0.9	4.7	0.9
R-3BLOCKS	0.819	0.225	0.722	0.211	4.5	0.8	4.3	0.7
R-CRIT	0.826	0.134	0.750	0.170	4.5	0.9	3.9	0.9

			Old patterns: Distortion level 5 (O5)					
T-SYST	0.854	0.206	0.854	0.185	4.5	0.9	4.6	0.8
T-SC-RAD	0.868	0.123	0.819	0.121	4.6	0.5	4.3	0.7
T-SC-CAT	0.903	0.106	0.813	0.139	4.6	0.9	4.1	1.1
R-3BLOCKS	0.799	0.216	0.611	0.266	4.4	0.8	3.8	0.9
R-CRIT	0.799	0.130	0.715	0.218	4.5	0.9	3.9	0.9
			Gappers: Distortion level 1 (G1)					
T-SYST	0.986	0.038	0.924	0.171	5.4	0.7	5.4	0.9
T-SC-RAD	0.972	0.086	0.993	0.028	5.6	0.5	5.7	0.4
T-SC-CAT	0.910	0.116	0.951	0.090	5.3	0.6	5.1	0.7
R-3BLOCKS	0.951	0.099	0.896	0.205	5.5	0.5	5.1	0.7
R-CRIT	0.931	0.090	0.931	0.090	4.9	0.7	5.0	0.7
			Gappers: Distortion level 2 (G2)					
T-SYST	0.944	0.122	0.917	0.170	5.2	0.9	5.2	0.9
T-SC-RAD	0.965	0.088	0.951	0.090	5.4	0.5	5.6	0.5
T-SC-CAT	0.951	0.081	0.882	0.137	5.1	0.8	5.1	0.8
R-3BLOCKS	0.903	0.202	0.875	0.202	5.0	0.8	4.7	1.0
R-CRIT	0.931	0.090	0.889	0.162	4.5	1.0	4.3	0.9
			Gappers: Distortion level 3 (G3)					
T-SYST	0.951	0.090	0.903	0.167	5.2	0.9	5.1	0.8
T-SC-RAD	0.951	0.090	0.938	0.081	5.5	0.6	5.4	0.5
T-SC-CAT	0.944	0.081	0.917	0.149	5.2	0.8	4.9	0.9
R-3BLOCKS	0.840	0.263	0.785	0.235	3.8	1.0	4.0	1.1
R-CRIT	0.792	0.134	0.785	0.174	4.0	1.1	3.6	1.1
			New exemplars: Distortion level low (N1)					
T-SYST	0.944	0.091	0.924	0.180	4.8	1.1	5.0	1.1
T-SC-RAD	0.924	0.105	0.944	0.099	5.0	0.8	5.2	0.7
T-SC-CAT	0.882	0.160	0.868	0.158	4.4	1.0	4.8	0.9
R-3BLOCKS	0.924	0.150	0.875	0.222	5.2	0.7	5.1	0.7
R-CRIT	0.951	0.090	0.931	0.151	4.8	0.8	4.7	0.8

APPENDIX A—*Continued*

	Classification				Recognition			
	Immediate		Delayed		Immediate		Delayed	
Condition	Mean	SD	Mean	SD	Mean	SD	Mean	SD
New exemplars: Distortion level medium (N2)								
T-SYST	0.604	0.327	0.750	0.255	2.2	0.8	2.6	1.0
T-SC-RAD	0.542	0.342	0.597	0.218	1.9	0.6	2.8	0.9
T-SC-CAT	0.507	0.287	0.576	0.267	2.3	0.8	2.7	1.2
R-3BLOCKS	0.764	0.292	0.653	0.289	3.1	0.8	3.1	1.0
R-CRIT	0.722	0.293	0.743	0.286	3.1	1.3	2.9	1.2
New exemplars: Distortion level high (N3)								
T-SYST	0.410	0.316	0.535	0.251	1.7	0.8	2.2	1.0
T-SC-RAD	0.347	0.069	0.382	0.233	1.5	0.4	1.8	0.6
T-SC-CAT	0.333	0.325	0.347	0.263	1.6	0.7	1.9	0.9
R-3BLOCKS	0.611	0.253	0.549	0.291	2.1	0.7	2.5	0.9
R-CRIT	0.590	0.283	0.535	0.227	2.3	0.9	2.4	0.9
Prototypes								
T-SYST	0.979	0.083	0.917	0.228	5.5	0.8	5.3	0.8
T-SC-RAD	0.958	0.114	0.938	0.134	5.6	0.6	5.5	0.6
T-SC-CAT	0.938	0.134	0.917	0.149	5.2	0.8	5.1	1.1
R-3BLOCKS	0.979	0.083	0.833	0.243	5.6	0.6	5.3	0.8
R-CRIT	0.958	0.014	0.938	0.134	5.5	0.6	5.3	0.7
Foils								
T-SYST	0.722	0.335	0.618	0.363	1.2	0.3	1.4	0.6
T-SC-RAD	0.722	0.354	0.632	0.336	1.3	0.3	1.5	0.4
T-SC-CAT	0.632	0.444	0.597	0.391	1.2	0.4	1.5	0.7
R-3BLOCKS	0.299	0.289	0.514	0.393	1.8	0.8	2.0	0.8
R-CRIT	0.403	0.410	0.542	0.376	1.8	0.8	2.0	0.7

APPENDIX B: CONSTRUCTION OF TEST ARRAYS FOR EXPERIMENT 2

Construction of the test array, which always contained three patterns (T5, T5′, T5″), involved three steps: (a) divide the distance between each point in the prototype and a high level distortion by 5, e.g., if the coordinates of one point of the prototype were (−1.36, 6.93) and the corresponding coordinate values of a high level distortion were (−5.93, 7.16), then the overall displacement would be (−4.57, 0.23); dividing by 5 would result in (−0.91, 0.05). These latter values were the increments for that point for each successive transformation from P to T5, e.g., T1 = (−1.36, 6.93) + (−.91, 0.05) = (−2.27, 6.98); (b) the increment for each pattern displacement/5 was reversed, as was one sign, and added to the coordinates for T4 to create one of the test array patterns, e.g., the displacement of (−.91, .05) becomes (.05, .91), which when added to T4, produces the corresponding coordinate values for T5, e.g., (−4.95, 8.04); and (c) the coordinates for T5″ were computed by reversing both signs of the (*x*,*y*) displacement for the first distortion, e.g., T5″ becomes T4 + (−.05, −.91) = (−5.05, 6.22). This procedure was followed for each of the nine points of each pattern. Every high level distortion, therefore, appeared with two distortions (T5′, T5″), where each pattern in the array was a minimal distortion of each other.

REFERENCES

Alexander, T. M., & Enns, J. T. (1988). Age changes in the boundaries of fuzzy categories. *Child Development,* **59,** 1372–1386.

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review,* **98,** 409–429.

Ashby, F. G., Boynton, G., & Lee, W. W. (1994a). Categorization response time with multidimensional stimuli. *Perception and Psychophysics,* **55,** 11–27.

Ashby, F. G., & Maddox, W. T. (1994a). A response time theory of separability and integrality in speeded classification. *Journal of Mathematical Psychology,* **38,** 423–466.

Ashby, F. G., & Maddox, W. T. (1994b). Relations between prototype, exemplar, and decision bound models of categorization. *Journal of Mathematical Psychology,* **37,** 372–400.

Ashby, F. G., Maddox, W. T., & Lee, W. W. (1994b). On the dangers of averaging across subjects when using multidimensional scaling or the similarity-choice model. *Psychological Science,* **5,** 144–151.

Biederman, I., & Gerhardstein, P. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance,* **19,** 1162–1182.

Biederman, I., & Shiffrar, M. M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual learning task. *Journal of Experimental Psychology,* **83,** 486–490.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology,* **11,** 177–220.

Brehmer, B. (1974). Hypotheses about the relations between scaled variables in the learning of probabilistic interference tasks. *Organizational Behavior and Human Performance,* **11,** 1–27.

Busemeyer, J. R., Byun, E., Delosh, E. L., & McDaniel, M. A. (1997). Learning functional relations based on experience and input-output pairs by humans and artificial neural networks. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts and categories* (pp. 408–437). Cambridge, MA: The MIT Press.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of the 'Eckart-Young' decomposition. *Psychometrika,* **41,** 439–463.

Cutting, J. E. (1978). Perceiving the geometry of age in a human face. *Perception and Psychophysics,* **24,** 566–568.

DeRosa, D. V., & Tkacz, S. (1976). Memory scanning for organized visual material. *Journal of Experimental Psychology: Human Learning and Memory,* **2,** 688–694.

Ebeling, W. (1991). Mutations and selection in evolutionary processes. In E. Mosekilde and L. Mosekilde (Eds.), *Complexity, chaos, and biological evolution*. New York: Plenum.

Estes, W. K. (1986). Array models for category learning. *Cognitive Psychology,* **18,** 500–549.

Finke, R. A., & Freyd, J. J. (1985). Transformation of visual memory induced by implied motions of pattern elements. *Journal of Experimental Psychology: Learning Memory and Cognition,* **11,** 780–794.

Freyd, J. J. (1987). Dynamic mental representations. *Psychological Review,* **94,** 427–438.

Freyd, J. J., & Johnson, J. Q. (1987). Probing the time course of representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **13,** 259–269.

Gelman, S. A., & Kremer, K. (1991). Understanding natural causes: Children's explanations of how objects and their properties originate. Child Development, **62,** 396–414.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General,* **123,** 178–200.

Goldstone, R. L., Steyvers, M., & Larimer, K. (1996). Categorical perception of novel dimensions. *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society* (pp. 243–248). Hillsdale. New Jersey: Erlbaum.

Gibson, E. J. (1969). *Principles of perceptual learning and development.* New York: Appleton–Century–Crofts.

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: Learning Memory and Cognition,* **13,** 87–103.

Grierer, A., & Meinhardt, H. (1972). A theory of biological pattern formation. *Kybernetik,* **12,** 30–39.

Hampton, J. A. (1987). Inheritance of attributes in natural concept conjunctions. *Memory and Cognition*, **15,** 55–71.

Hayes-Roth, B., & Hayes-Roth, F. (1977). Concept learning and the recognition and classification of exemplars. *Journal of Verbal Learning and Verbal Behavior,* **16,** 321–338.

Heit, E., & Rubinstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning Memory and Cognition,* **20,** 411–422.

Hintzman, D. L. (1986). Schema abstraction in a multiple-trace memory model. *Psychological Review,* **93,** 411–428.

Homa, D. (1978). Abstraction of ill-defined form. *Journal of Experimental Psychology: Human Learning and Memory,* **4,** 407–416.

Homa, D. (1984). On the nature of categories. In G. H. Bower (Ed.), *The psychology of learn-*

ing and motivation: Advances in research and theory (Vol. 18, pp. 49–94). San Diego, CA: Academic Press.

Homa, D., Dunbar, S., & Nohre, L. (1991). Instance frequency, categorization, and the modulating effect of experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **17,** 444–458.

Homa, D., Rhoads, D., & Chambliss, D. (1979). Evolution of conceptual structure. *Journal of Experimental Psychology: Human Learning and Memory,* **5,** 11–23.

Homa, D., & Vosburgh, R. (1976). Category breadth and the abstraction of category information. *Journal of Experimental Psychology,* **2,** 322–330.

Hull, C. L. (1920). Quantitative aspects of the evolution of concepts. *Psychological Monographs,* **28** (1, Whole No. 123).

Jonides, J. (1981). Voluntary versus automatic control over the mind's eye's movement. In J. Long & A. Baddely (Eds.), *Attention and performance* (pp. 187–203). Hillsdale, NJ: Erlbaum.

Keil, F. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.

Knapp, A. G., & Anderson, J. A. (1984). Theory of categorization based on distributed memory storage. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **13,** 87–108.

Koh, K., & Meyer, D. E. (1991). Function learning: Induction of continuous stimulus-response relations. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **17,** 811–836.

Kruschke, J. K. (1992). ALCOVE: An exemplar based model of category learning. *Psychological Review,* **99,** 22–44.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychological Review,* **29,** 1–27.

Kruskal, J. B., Young, F. W., & Seery, J. B. (1973). *How to use KYST, a very flexible program to do multidimensional scaling and unfolding*. Murray Hill, NJ: Bell Laboratories.

Kuhl, P. K. (1991). Human adults and human infants show a ''perceptual magnet affect'' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics,* **50,** 93–107.

Lakoff, G. L. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: Univ. of Chicago Press.

Lamberts, K. (1998). The time course of categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **24,** 695–711.

Lawson, R., & Humphreys, G. (1996). View specificity in object processing: Evidence from picture matching. *Journal of Experimental Psychology: Human Perception and Performance,* **22,** 395–416.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception and Psychophysics,* **53,** 49–70.

McLean, J. P., & Shulman, G. L. (1978). On the construction and maintenance of expectancies. *Quarterly Journal of Experimental Psychology,* **30,** 441–454.

Meakin, P. (1986). A new model for biological pattern formation. *Journal of Theoretical Biology,* **118,** 101–113.

Medin, D. L., & Bettger, J. G. (1994). Presentation order and recognition of categorically related examples. *Psychonomic Bulletin and Review,* **1,** 250–254.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review,* **100,** 254–278.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review,* **85,** 207–238.

Medin, D. L., Wattenmaker, W. D., & Hampson, S. E. (1987). Family resemblance, conceptual cohesiveness, and category construction. *Cognitive Psychology,* **19,** 242–279.

Muller, J., Rambaek, J. P., Hovig, C., & Hovig, T. (1991). Multifractal analysis of morphological patterns in normal and malignant human tissues. In E. Mosekilde and L. Mosekilde (Eds.), *Complexity, chaos, and biological evolution*. New York: Plenum.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review,* **90,** 339–363.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, **115,** 39–57.

Nosofsky, R. M. (1987). Attention and learning processes in the identification-categorization relationship. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **13,** 87–108.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **14,** 54–65.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review,* **104,** 266–300.

Olivier, D. C. CONGRU. [computer program]

Phillips, J. B. (1975). *Development of vertebrae anatomy*. London: Mosby.

Pinker, S., Choate, P. A., & Finke, R. A. (1984). Mental extrapolation in patterns constructed from memory. *Memory and Cognition,* **12,** 207–218.

Pittenger, J. B., & Shaw, R. E. (1975). Aging faces as viscal-elastic events: Implications for a theory of nonrigid shape perception. *Journal of Experimental Psychology: Human Perception and Performance,* **1,** 374–378.

Pittenger, J. B., Shaw, R. E., & Mark, L. S. (1979). Perceptual information for the age level of faces as a higher order invariant of growth. *Journal of Experimental Psychology: Human Perception and Performance,* **5,** 478–493.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, **77,** 353–363.

Posner, M. I., Goldsmith, R., & Welton, K. E. (1967). Perceived distance and the classification of distorted patterns. *Journal of Experimental Psychology,* **73,** 28–38.

Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. In P. M. A. Rabbit & S. Dornic (Eds.), *Attention and performance V* (pp. 669–682). San Diego, CA: Academic Press.

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21–49). New York: Cambridge Univ. Press.

Rips, L. J., & Collins, A. (1993). Categories and resemblance. *Journal of Experimental Psychology: General*, **122,** 468–486.

Rosch, E. (1975a). The nature of mental codes for color categories. *Journal of Experimental Psychology: Human Perception and Performance,* **1,** 303–322.

Rosch, E. (1975b). Cognitive representations in semantic categories. *Journal of Experimental Psychology: General*, **104,** 192–233.

Ross, B. H., Perkins, S. J., & Tenpenny, P. L. (1990). Reminding-based category learning. *Cognitive Psychology,* **22,** 460–492.

Samarapungavan, A. (1992). Children’s judgments in theory choice tasks: Scientific rationality in childhood. *Cognition*, **45,** 1–32.

Seamon, J. G. (1982). Dynamic facial recognition: Examination of a natural phenomenon. *American Journal of Psychology,* **95,** 363–381.

Seamon, J. G., Stoltz, J. A., Bass, D. A., & Chatinover, A. I. (1978). Recognition of facial features in immediate memory. *Bulletin of the Psychonomic Society,* **12,** 231–234.

Shepard, R. N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function: I. *Psychometrika,* **27,** 125–140.

Shin, H. J., & Nosofsky, R. M. (1992). Similarity-scaling studies of dot-pattern classification and recognition. *Journal of Experimental Psychology: General,* **121,** 278–304.

Sneath, P. H. A., & Sokal, R. R. (1973). *Numerical taxonomy*. San Francisco, CA: Freeman.

Stone, J. V. (1998). Object recognition using spatiotemporal signatures. *Vision Research*, **38,** 947–951.

Tversky, A. (1977). Features of similarity. *Psychological Review,* **84,** 327–352.

Ullman, S. (1979). *The interpretation of visual motion.* Cambridge, MA: MIT Press.

Verfaillie, K., & D'Ydewalle, G. (1991). Representational momentum and event course anticipation in the perception of implied periodical motions. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* **17,** 302–313.

Wood, D. S. (1983). A phenetic assessment of the ciconiidae (aves) using skeletal morphology. *Annals of Carnegie Museum*, **52,** 79–112.

Accepted April 6, 1999