Categorization and stimulus structure

Steven Ritter
Department of Psychology
Carnegie Mellon University
Pittsburgh, Pa. 15213
sritter@psy.cmu.edu

Abstract

Concept discovery experiments have yielded theories that work well for simple, rule-governed categories. They appear less applicable to richly structured natural categories, however. This paper explores the possibility that a complex but structured environment provides more opportunities for learning than the early theories allowed. Specifically, category structure may aid in learning in two ways: correlated attributes may act jointly, rather than individually, and natural structure may allow more efficient cue sampling. An experiment is presented which suggests that each of these advantages may be found for natural categories. The results call into question independent sampling assumptions inherent in many concept learning theories and are consistent with the idea that correlated attributes act jointly. In order to model natural category learning, modifications to existing models are suggested.

Introduction

Thus, commencing our investigation by a careful survey of any one bone by itself, a person who is sufficiently master of the laws of organic structure, may, as it were, reconstruct the whole animal to which that bone belonged.

The quote, written by the naturalist Georges Cuvier in 1812, expresses confidence that the laws of nature so determine the structure of natural kinds that we should be able to deduce the entire form of an animal from only the smallest part. Today, few would argue that natural structure is so strongly determined, but fewer still would argue that the features of a natural object are assembled without regard to those already present. Rather, the laws of nature constrain the co-occurrence of features. Although wings and hollow bones do not necessarily go together, we are likely to find one, having found the other.

Psychologists are only beginning to examine the implications of natural structure for learning. In concept discovery tasks, the issue was usually avoided. A subject might be asked to learn the concept "large red square" or a "group of two items." The dimensions on which concepts varied were selected to be obvious to the subject and convenient for the experimenter to manipulate. There was little concern about whether dimensions were structured or rules defined as they would be in natural kinds. Such experiments taught us quite a bit about learning. For example, they suggested that subjects sampled from a hypothesis space, changing their working hypothesis only when they made an error (e.g., Restle, 1963; Trabasso and Bower, 1968).

Rosch and her colleagues (e.g., Rosch, et al.., 1977) revived interest in the structure of natural kinds. Their work showed that natural categories are not at all like the assemblage of features used to represent a concept in concept discovery experiments. Instead, natural kinds appear to be structured around "family resemblances." Family resemblance categories consist of a large number of highly inter-correlated features. To a large extent, these correlations are not arbitrary; rather, they are the result of natural laws.

This research was supported in part by a graduate student fellowship from the National Science Foundation.

Findings about natural category structure call into question the relevance of concept discovery methods for learning natural kinds because the conditions found in natural categories appear to be precisely those which make concept discovery difficult. Concept discovery experiments turn out to be intractable in all but the simplest versions. If disjunctive rules (e.g., "red or square") are allowed, or if the rule comes from a non-standard space (such as when the correct response depends on a past pattern of responses), subjects may search indefinitely without finding the correct solution (Levine, 1975). If the number of irrelevant attributes is high, subjects will be unable to learn in a reasonable amount of time (Bourne and Haygood, 1959).

Correlations among relevant attributes may help, but only a little. Trabasso and Bower (1968) found that, when stimuli contained features perfectly correlated with each other and with the category label, subjects tended to categorize by either one cue or the other. Only rarely did a subject notice the relationship between both cues and the category label. More importantly, these experiments showed that redundant relevant cues tend to compete for attention so that, although the overall probability of correctly learning increases as the number of relevant dimensions increase, the benefit is only due to having more predictive features. Subjects sample only a small set of features on each trial, so the benefit diminishes as the number of correlated attributes increases.

A central question, then, is how natural categories are learnable at all. Although children may take some time to learn the difference between a dog and a cat, the actual number of exposures to such animals over that time is typically much smaller than the number of exemplars viewed in a concept discovery experiment.

Billman and Heit (1988) present one part of the answer. In their model, each trial provides an opportunity to predict one feature based on another. The prediction uses previously observed contingencies between the predictor value and values of the predicted feature. The choice of predictor and predicted features is governed by salience, which increases for both the predictor and predicted feature when a correct prediction is made. This relationship sets up a feedback loop in which correlations lead to correct predictions, which increase the salience of their constituent features. This results in a greater chance of noticing correlations involving these features, so the cycle continues. Thus, according to this model, the correlation between "has feathers" and "bird" is learned more easily because there is a correlation between "has feathers" and "flies." Billman (1989) provides experimental support for the model.

Correlated attributes can only be part of the answer, however. Natural categories have too many potential correlations to be effectively searched without some guidance. This study aims to provide further evidence for mutual support of correlated attributes and for search guided by structural constraints. The underlying belief is that natural categories may be learnable because their rich structure allows us to make assumptions about which features are likely to go together. While it seems logical to assume that we have developed a categorization system that can take advantage of natural structure, it has been difficult to demonstrate any such advantage (Medin, Wattenmaker and Hampson, 1987; Wattenmaker, et al., 1986; but see Malt and Smith, 1984).

In the current experiment, the stimuli, although artificial, have a rich correlational structure. Redundancy of cues ensures that a stimulus can be unambiguously classified, even if all relevant cues are not present. This is equivalent to the natural situation, where we typically have access to only a few of the many cues we might want to use in identifying something. We make two predictions. First, we expect a well-structured category to be easier to learn than an ill-structured category, even if each provides equivalent information. Second, we expect to replicate Billman's finding that correlated attributes, when embedded in a logical structure, will aid learning beyond the level that would be expected if the cues were acting independently.

Method

Subjects

Subjects were 54 undergraduates taking an introductory psychology class. Participation in the experiment partially fulfilled a course requirement.

Stimuli

Stimuli consisted of pictures of boats, each containing a maximum of five parts: the mainsail, the jib sail, the flag, the rudder and the centerboard. Each of these parts could vary in color and sometimes shape. The body and mast of the boat were identical in all stimuli and were always present (see Figure 1). Any of the variable parts could be missing. The probability that a particular part was present in a stimulus (i.e., the part's availability) is given in Table 1.

Three conditions (control, well-structured and ill-structured) were run. The conditions differed in the composition of stimuli. In the "well-structured" condition, each color corresponded to a shape. For example, all blue flags in this condition were short and wide, while all pink flags were long and thin. In the "ill-structured" condition, the color of a part was independent of its shape. In the control, only color changed.

Stimuli were assigned to two classes, called "gemmer" and "brice." Stimuli could be classified correctly using either color or shape of the parts of the boat. In all conditions, boats could be classified according to the rule: "If the flag is blue, the boat is a gemmer. If the flag is missing and the main sail is blue, the boat is a gemmer. If both the flag and mainsail are missing and the remaining sail is blue, the boat is a gemmer." The rule can be thought of as a hierarchy. The flag is the most important part of

	Flag	Mainsail	Jib	Rudder	Center-board
Availability	65	67	85	67	65
Reliability	100	65	57	35	10

Table 1: Availability and reliability of color for stimuli in the learning phase. In the well-structured group, availabilities and reliabilities for shape are identical to those for color. For shapes in the ill-structured group, values for flag and mainsail apply to the centerboard and rudder (respectively). Availability refers to the percentage of time the part is present. Reliability is the percentage of time the part's color (or shape) predicts the category.

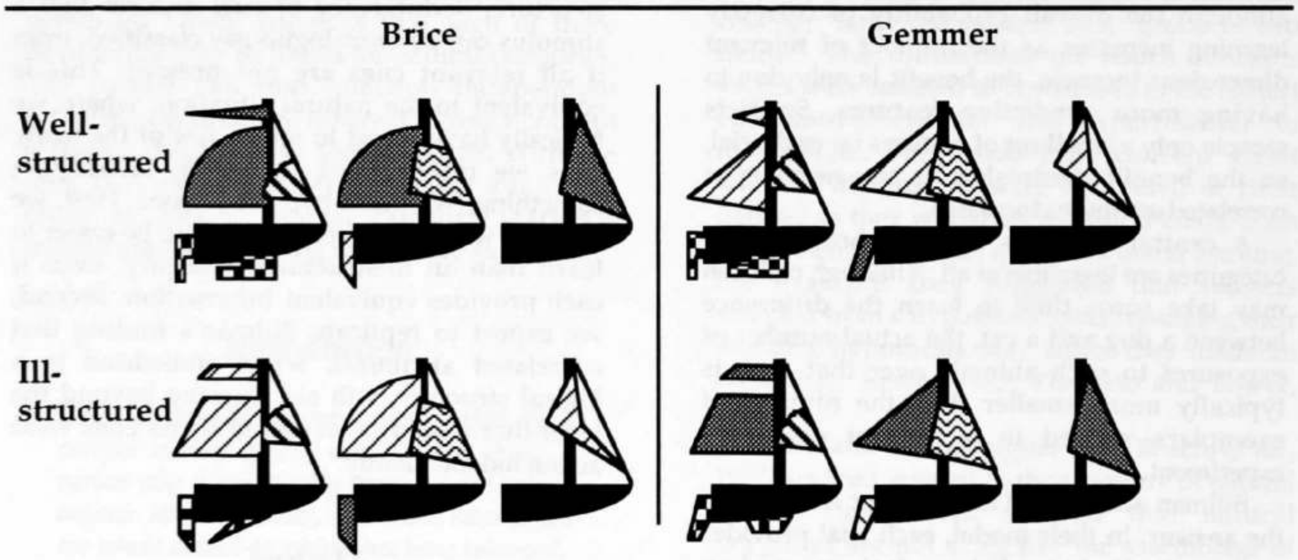

Figure 1: Examples of stimuli used in the experiment. Patterns represent colors used in the experiment. Corresponding boats in A and B have the same coloring. Either set can be categorized by the rule: "If the flag is shaded, it is a brice. If there is no flag and the mainsail is shaded, it is a brice. If there is neither a flag nor a mainsail and the jib is shaded, it is a brice." The ill-structured group can also be classified by a rule of the form: "If the centerboard is triangular, it is a gemmer. If there is no centerboard and the rudder points down, it is a gemmer. If there is no centerboard or rudder and the jib is a concave polygon, it is a gemmer." A similar shape rule applies to the well-structured group.

the boat, followed by the mainsail and the jib sail. In addition, each kind of boat has a color: blue for gemmer and pink for brice. For each stimulus, the topmost part in the hierarchy was guaranteed to be predictive. As a result, flags were always either pink or blue, and, for boats without flags, mainsails were always either pink or blue. In the well-structured and ill-structured conditions, subjects could additionally categorize exemplars by a rule referring to shape, rather than color. In the well-structured condition, the shape hierarchy was the same as the color hierarchy (flag, then mainsail, then jib). In the ill-structured condition, the color hierarchy was flag, then mainsail, then jib; but the shape hierarchy was centerboard, then rudder, then jib. Stimuli in the well-structured and ill-structured conditions were matched such that equivalent stimuli in the two conditions provided both color and shape cues at the same level of the hierarchy. Stimuli in the control condition were matched on the color hierarchy.

Further constraints were put on the stimuli to provide a richer category structure. Four of the five parts had a greater than chance probability of being the color or shape predictive of the category. The reliability of each part of the boat is given in Table 1.

Procedure

The experiment was run on a color Macintosh computer. The experimental session consisted of two stages: training and test. For each trial in the training stage, a stimulus was presented along with the choices "Brice" and "Gemmer." Subjects gave their responses by clicking on the appropriate answer. Feedback followed immediately in the form of "Yes (no) this is a gemmer (brice)." Subjects could study the stimulus along with the feedback and initiate the next trial when they were ready. The training stage continued for 60 trials.

In the test stage, subjects responded to stimuli identical to those in the training stage as well as to stimuli that omitted either the color or shape cues. "Shape-only" stimuli were drawn in gray. "Color-only" stimuli displayed colored circles in place of the studied shapes. Control subjects were not given "shape-only" stimuli. Fourteen stimuli of each kind were presented, for a total of 28 test trials in the control group and 42 in each of the other groups. Test stimuli were constructed so that the color-only, shape-only and color-and-shape sets depended on each level of the hierarchy to the same extent.

The procedure for test stimuli was the same as for training stimuli, except that feedback was not given. Instructions for the test stage were not given until the training stage was completed, so subjects were not biased to look at both color and shape. Responses and reaction times were recorded in both stages, but the instructions emphasized correct responding only.

Cue Learned	Well-structured	Ill-structured
Color Only	8	2
Shape Only	1	5
Color and Shape	5	2
Neither	4	9

Table 2: Number of subjects learning to respond to each (or both) cues, by group. A subject was considered to be responding to a cue if the subject correctly answered at least 10 of 14 questions providing only that cue.

Results

A comparison of the well-structured and ill-structured groups supports our hypothesis about stimulus structure. We may classify subjects as having learned the color cue, the shape cue or both cues, depending on their performance on the testing phase. Subjects correctly answering at least 10 of the 14 color-only questions and less than 10 of the 14 shape-only questions were classified as "color learners." Those correctly answering 10 of 14 shape-only questions and less than 10 color-only questions were classified as "shape learners." Those correctly answering more than 10 in each category were classified as "color and shape learners."[1]

Table 2 shows how subjects from the well-structured and ill-structured groups were

[1]The binomial probability of a subject with no knowledge of either cue falling into either the color-only or shape-only category is .08. The probability of such a subject falling into the "color and shape" category is .01.

classified. Eighty percent of color learners were in the well-structured group. Although the majority of shape learners were in the ill-structured group, this statistic may be misleading. All but one shape learner in the well-structured group also responded to color, thus falling in the "color and shape" group. In total, 13 subjects in the well-structured group and 4 in the ill-structured group learned to respond to color. Six subjects in the well-structured group and 7 in the ill-structured group learned to respond to shape.

There are other indications that the well-structured group found the task somewhat easier. The amount of time taken to learn (in the training stage) was measured by a criterion of trials to second-to-last error (this statistic was more resistant to careless errors than trials to last error). Subjects in the well-structured group reached criterion in an average of 40.0 trials, while those in the ill-structured group took 48.3 trials. An analysis of variance showed the difference to approach significance, $F(1,30)=3.64$, $p<.07$.

Stronger support for the hypothesis comes from the test phase (see Figure 2). A 2 (group) x 3 (learning stimulus order) analysis of variance of percent correct on the color-only questions shows a difference between groups, $F(1,30)=6.17$, $p<.02$. Reaction times tell a similar story. These times were subjected to a log transformation and then analyzed in a 2 (group) x 3 (learning stimulus order) x 14 (question) analysis of variance. The results show a significant difference between groups, $F(1,311) = 10.10$, $P<.01$.

Analyses of shape-only and color-and-shape questions were less conclusive. On color-and-shape questions, the pattern of results (both in percentage correct and reaction time) is the same as for color only, but the differences do not reach significance, $F(1,30)= 0.55$, $p>.46$. Shape-only questions show some advantage for the ill-structured group, but the results do not approach significance, $F(1,30) =1.18$, $p>.28$.

A comparison of the well-structured and control groups speaks to our second prediction, that correlated attributes contribute beyond their individual influences. If correlated attributes support each other, we would expect the well-structured group to be able to answer color-only and color-and-shape questions more easily than the control group. This result would obtain if the correlation between shape and color (found in the well-structured stimuli) helped subjects discover the correlation between color and category label. The experimental results are not strong with respect to this hypothesis. Subjects in the well-structured group learned faster than those in the control (by trials to second-to-last error, $F(1,30) = 4.69$, $p<.05$), but this could have been due to the independent influence of shape. Both reaction times and percent correct scores show an advantage for the well-structured group (see Figure 2), but the differences are not statistically significant.

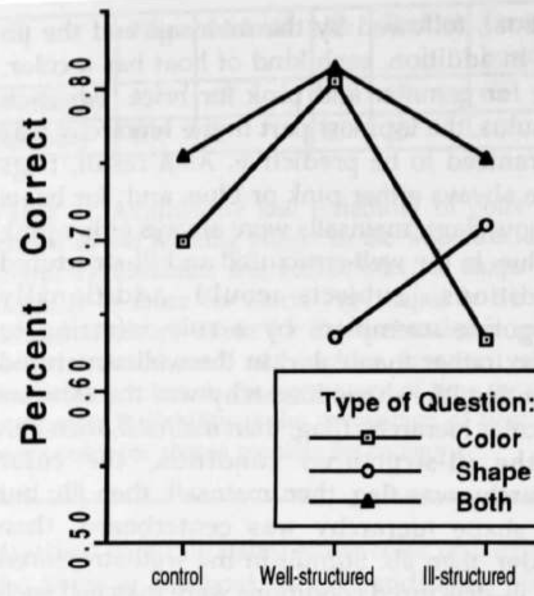

Figure 2: Performance by question type

Discussion

The data reflect on two aspects of categorization models. The first is whether correlated attributes contribute individually, as in the Trabasso and Bower model or jointly, as in the Billman and Heit model. While the data show some advantage for the well-structured group over the control, the results do not approach significance. This should not be taken as a failure to replicate Billman and Heit's model, since that model predicts much stronger effects in situations where no feedback is given.

The second issue is whether stimulus structure influences cue sampling. Since the well-structured and ill-structured conditions provide the same number of cues, both the Trabasso and Bower model and the Billman and Heit model predict that there would be no difference between the well-structured and ill-structured groups. If cue-sampling is sensitive to stimulus structure, however, we might expect an advantage for the well-structured condition. In this case, the fact that flag color was relevant might direct us to notice that flag shape is also relevant (in the well-structured condition). The experimental results provide some support for this kind of model. Subjects in the well-structured condition were better able to answer questions about the color cue than subjects in the ill-structured condition, and they discovered the categorization rule faster than other subjects. In addition, these subjects were able to make judgments about color-only stimuli faster than subjects in the ill-structured group. The finding that the advantage for the well-structured group reversed (albeit non-significantly) for the shape-only stimuli is problematic for this explanation, however.

The experiment presented here was designed to examine whether structural properties of complex, real-world objects aid in learning categories. The results suggest that structural properties play a role by influencing the order in which cues are sampled. Clapper and Bower (1991) present a model in which correlated attributes in a category influence the sampling of cues in new exemplars. Pazanni (1991) presents a model in which domain-specific biases affect the order of hypothesis search. The way in which these different influences on search interact is a topic for future research.

Acknowledgments

I thank Brian MacWhinney for his guidance in all stages of this work.

References

Billman, D. (1989). Systems of correlations in rule and category learning: Use of structured input in learning syntactic categories. *Language and Cognitive Processes*, 4, 127-155.

Billman, D. and Heit, E. (1988). Observational learning from internal feedback: A simulation of an adaptive learning method. *Cognitive Science*, 12, 587-625.

Bourne, L. E., Jr. and Haygood, R. C. (1959). The role of stimulus redundancy in concept identification. *Journal of Experimental Psychology*, 58, 232-238.

Clapper, J. and Bower, G. H. (1991) Learning and applying category knowledge in unsupervised domains in G. H. Bower (Ed.), *The Psychology of Learning and Motivation, Volume 27*. (pp. 65-108). New York: Academic Press.

Cuvier, Georges (1812). *Discours préliminaire*. Quoted in Gould, S. J. (1990, June). Everlasting legends. *Natural History*, 12-17.

Levine, M. A. (1975). *A cognitive theory of learning: Research on hypothesis testing*. Hillsdale, NJ: Erlbaum.

Malt, B. C. and Smith, E. E. (1984). Correlated properties in natural categories. *Journal of Verbal Learning and Verbal Behavior*, 23, 250-269.

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

Pazzani, M. J. Influence of prior knowledge on concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 17, 416-432.

Restle, F. (1962). The selection of strategies in cue learning. *Psychological Review*, 69, 11-19.

Rosch, E., Mervis, C., Gray, W., Johnson, D. and Boyes-Braem, P. (1976), Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.

Trabasso, T. and Bower, G. H. (1968). Attention in learning: theory and research. New York: Wiley.

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D. and Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties and concept naturalness. *Cognitive Psychology*, 18, 158-194.