Formation of ill-defined concepts as a function of category size and category exposure

MARY JANE DiNARDO and THOMAS C. TOPPINO
Villanova University, Villanova, Pennsylvania

Three experiments investigated the formation of ill-defined concepts using dot pattern stimuli. Previous research had varied category size (the number of different category exemplars presented during training) and found that subsequent transfer to new exemplars improved as a function of increasing category size. However, the category size effect could have been due partially to a bias to classify new stimuli into larger categories and/or to the amount of exposure to the category, a factor that was confounded with category size. Experiment 1 removed the source of potential response bias and still obtained a category size effect, although the latter variable remained confounded with category exposure. Experiments 2 and 3 varied category size and category exposure independently. Results suggested that category exposure may have contributed to findings that had previously been attributed to category size.

One of the most fundamental of our cognitive abilities is our capacity to classify or categorize stimulus information even when the requirements for category membership seem to be ill-defined. It is usually assumed that this ability depends upon and reflects underlying concepts that we have acquired through experience. Two important research questions are: (1) What is acquired? That is, what is the nature of the information or representation on which classifications are based? (2) What are the factors and/or processes that affect the formation of these representations? Of course, these two questions are necessarily interdependent because classification of a stimulus depends on what has been learned, which in turn may depend on how learning takes place.

One factor that seems to affect the formation of ill-defined concepts is category size, which is defined as the number of different exemplars of a category that a subject experiences in the course of its acquisition (e.g., Homa & Chambliss, 1975; Homa, Cross, Cornell, Goldman, & Shwartz, 1973) For example, Homa et al. (1973) trained subjects to classify 18 visually presented dot patterns into three categories represented by three, six, and nine training patterns, respectively. Training continued until subjects met a criterion of two consecutive trials in which all 18 training stimuli were classified correctly. Subsequently, a transfer test was administered in which subjects attempted to classify some old training patterns from each category as well as new patterns from each category, including prototypes representing the central tendency of each category. Results indicated that

transfer performance on new patterns improved as a function of category size. That is, new patterns were classified more accurately if they were members of categories that had been represented by more patterns during training.

This finding, the category size effect, has usually been taken as evidence that, in transfer, classification of new patterns is based on similarity to subjective prototypes of each category that were abstracted during training. These prototypes may be described theoretically in a number of ways, such as the average of the training patterns in each category or as a set of features that occur with modal frequencies in each category (Neumann, 1974; Posner, 1969; Reed, 1972). But, in any case, overall transfer to new patterns of a given category should depend, in part, on how well the subjective prototype approximates the true central tendency of the category. This approximation should be closer and transfer performance should be better when the subjective prototype is abstracted from a larger set of training stimuli.

Although the above interpretation of the category size effect seems intuitively compelling, there are other interpretations. For example, Medin and Schaffer (1978) proposed that subjects store representations of the individual training patterns and base transfer classifications on similarity to one or more of the training patterns. According to this view, increasing category size will increase the probability that a new transfer pattern will be similar to at least one training pattern from the appropriate category.

In the present research, we investigated two potential problems with the category size effect itself. First, previous research had employed within-subjects designs so that subjects simultaneously learned several categories of different sizes. Such within-subject procedures could

have created a bias to classify new patterns in the transfer task into the larger categories. Although statistical corrections have been used previously (e.g., Homa et al., 1973), our Experiment 1 eliminated the potential source of bias completely by using a between-subjects design. The second problem is that, in previous studies, category size has been confounded with total exposure to the categories. To illustrate, if, on each trial, subjects are exposed to three instances from one category and nine from another, they are exposed to one category three times and to the other nine times. If, however, each stimulus in the three-instance category were presented three times per trial, subjects would be exposed to each category nine times on each trial although category size or number of different instances would still vary. Our Experiment 2 held category exposure constant while varying category size. Experiment 3 held category size constant while varying category exposure.

The possibility that the category size effect may depend entirely or in part on category exposure presents difficulties both for theories that assume transfer to be based on similarity to an abstracted prototype and for theories that assume transfer to be based on similarity to stored training patterns. Both kinds of theories clearly predict a category size effect, but, at present, neither kind of theory incorporates a mechanism that would readily explain why transfer performance should be facilitated by increasing category exposure beyond the point required to learn the individual training patterns. That is, further experience or exposure to the same set of training patterns would neither increase the match between a subjective prototype and the true central tendency of a category nor increase the probability that a new transfer pattern would be similar to one of the training patterns.

EXPERIMENT 1

Method

Stimuli. The stimuli consisted of three categories of random dot patterns constructed using Posner, Goldsmith, and Welton's (1967) procedures. Briefly, a prototype for each category was constructed by placing nine dots randomly in the center 30 x 30 field of a 50 x 50 field on 10-lines/in. graph paper. To create other examplars of each category, the prototype was distorted by moving its dots according to a probabilistic decision rule. For each category, seven patterns were generated at each of three levels of distortion: low, medium, and high. These corresponded, respectively, to the 3.0-, 5.0-, and 7.7-bit distortions of Posner et al. (1967). For presentation during the experiment, each dot pattern was transferred to a 5 x 5 in. unlined index card, using a diameter of 2 cm for the dots.

Subjects and Design. The subjects were 40 introductory psychology students, who were assigned randomly in blocks of 4 to four groups that differed with regard to category size. That is, depending on the group, training was conducted on 3, 6, 9, or 12 patterns per category.

Procedure. The subjects participated individually. They were told that they would be shown patterns of dots, each belonging to one of three categories, called "A," "B," and "C." Their task would be to learn to classify each dot pattern correctly into its appropriate category. The subjects were told that all patterns

belonging to a particular category would bear some similarity to one another and that it would be easier to learn the patterns by learning these similarities. Noting the differences between patterns belonging to different categories would also help to distinguish the patterns belonging to different categories.

Depending on whether each category was represented by 3, 6, 9, or 12 training patterns, each training trial involved the presentation of 9, 18, 27, or 36 stimuli, respectively. However, regardless of category size, low-, medium-, and high-distortion patterns were always represented equally among the training stimuli of each category. Training patterns were presented successively, and the order of presentation was randomized for each subject on each trial by shuffling the stimulus cards.

For the first two trials, each pattern was presented for 5 sec, and the subjects were told its category membership. From the third trial to the end of training, presentation was self-paced. The subjects classified each pattern as it was presented and then received corrective feedback. Training continued until the subjects had attained a criterion of one errorless trial.

Immediately following training, the subjects received a single transfer trial. For each category, transfer stimuli included one old training stimulus at each distortion level, three new (previously unseen) stimuli at each distortion level, and the prototype. Stimuli were presented successively in random order, and the subjects classified each one without receiving feedback. Again, the presentation was self-paced.

Results

The mean number of training trials to criterion for the 3-, 6-, 9-, and 12-instance-per-category conditions were 9.3, 9.7, 11.0, and 13.4, respectively. Thus, trials to criterion increased as a function of category size. However, primary interest centered on transfer performance. A three-way analysis of variance involving category size as a between-subjects factor and distortion level and categories (A, B, C) as within-subjects factors was performed separately on the error data for old instances and the error data for new instances. For old members, no effects were reliable. For new members, Table 1 presents the mean proportion of transfer errors as a function of category size and distortion level. The proportion of transfer errors on new instances increased

Table 1
Mean Proportion of Errors on New Instances as a Function of
Distortion Level and Condition for Experiments 1, 2, and 3

Condition	Prototype	Low	Medium	High	Total
	Distortion Level				
Experiment 1					
3 Patterns per Category	.167	.243	.199	.534	.286
6 Patterns per Category	.133	.176	.209	.374	.223
9 Patterns per Category	.000	.143	.187	.454	.196
12 Patterns per Category	.000	.077	.099	.342	.130
Experiment 2					
3 Patterns per Category	.000	.133	.165	.532	.205
	(.073)	(.098)	(.257)	(.368)	(.199)
9 Patterns per Category	.000	.077	.143	.566	.197
	(.006)	(.092)	(.221)	(.399)	(.180)
Experiment 3					
Criterion	.175	.135	.174	.515	.249
Criterion Plus 8 Trials	.035	.098	.145	.444	.181

Note–Data in parentheses are from a replication of Experiment 2.

significantly as a function of distortion level [F(3,108) = 60.86, p < .01] and decreased significantly as a function of category size [F(3,36) = 3.43, p < .05].

A significant category effect [F(2,72) = 7.68, p < .01] was due to the fact that Category B produced fewer errors than either Category A or Category C. Finally, there was a significant three-way interaction of size x distortion level x category [F(18,216) = 2.09, p < .01]. The major basis for the interaction involved high distortions in Category C, for which subjects learning 6 items per category performed significantly better than those learning 3, 9, or 12 items per category. This qualification to the main effect of size is unexplained, but it seemed most likely that the high distortions that had been probabilistically generated for Category C were unusual in some way. Therefore, the patterns in Category C were replaced by a completely new set of stimuli for use in Experiments 2 and 3.

Discussion

The results of this experiment indicate that the category size effect can be obtained using a between-subjects design that eliminates any bias to classify new patterns into larger categories during transfer. However, as in previous studies, smaller categories received less category exposure during training than did larger categories. The mean number of times that the subjects were exposed to a stimulus from each category ranged from 27.9 for the 3-pattern-per-category group to 160.8 for the 12-pattern-per-category group.

EXPERIMENT 2

Method

Subjects and Design. Twenty introductory psychology students were assigned randomly in blocks of two to two category size groups that differed in whether subjects were trained on three or nine different training patterns per category. The data of two additional subjects who were originally assigned to the nine-instance-per-category condition were discarded because they failed to reach criterion by the 17th trial.

Stimuli and Procedure. The stimuli and procedures were essentially the same as those in Experiment 1. However, both groups received 17 training trials before the transfer test was administered. And each pattern in the three-pattern-per-category condition was presented three times on each trial. In this way, both groups were exposed to each category nine times on each trial so that total category exposure was equated between groups while category size or the number of different training patterns still varied.

Results and Discussion

Except for two subjects who were replaced, all subjects in both groups attained a criterion of one errorless training trial on or before the 17th training trial. The error data for old and new instances on the transfer test were submitted to separate analyses of variance involving category size as a between-subjects factor and distortion level and category as within-subjects factors.

The subjects classified old stimuli almost perfectly in the transfer test, and the analysis of variance yielded no reliable effects. The mean proportion of errors for new transfer patterns is presented in Table 1. Analyses indicated that transfer errors increased as a function of dis-

tortion level [F(3,54) = 82.09, p < .01]. There was also a significant category x distortion level interaction [F(6,108) = 2.43, p < .05], which was due to a greater proportion of errors on medium-distortion members of Category A than on medium-distortion members of Categories B and C. However, the most interesting result was that the effect of category size did not even approach significance [F(1,18) < 1.00].

These findings were subsequently replicated in an experiment involving 36 subjects (18 per group). Procedures were virtually identical, with the following major exceptions. First, stimuli were presented on a CRT, and a computer generated categories independently for each subject. Thus, results could not be attributed to the fortuitous use of a particular set of categories or a particular sample of stimuli from those categories. Second, because low and medium distortions had proved to be similar in our original experiments, medium distortions were changed so that they corresponded to the 6-bit distortions of Posner et al. (1967). Third, the subjects were given 15, rather than 17, training trials. The results are presented in parentheses in Table 1. The effect of distortion level was significant [F(3,102) = 43.00, p < .01], with all levels except the prototype and low distortions being reliably different from one another according to Tukey's (a) test, p < .01. However, the category size again failed to approach significance [F(1,34) < 1.00]. This was true even though there should have been a subject-selection bias in favor of the nine-instance-per-category condition because 5 subjects originally assigned to this condition failed to reach criterion in 15 trials and had to be replaced.

EXPERIMENT 3

Method

The subjects were 38 introductory psychology students, who were assigned randomly in blocks of 2 to two groups that differed in category exposure. The subjects in both groups were trained on three patterns per category, but one group was trained to a criterion of one errorless trial, whereas the other group was trained to that criterion plus eight additional trials. Otherwise, stimuli and procedures were exactly the same as they had been in the three-instance-per-category condition of Experiment 2.

Results

The groups did not differ in training trials to criterion. The error data for old and new instances on the transfer test were submitted to separate analyses of variance involving exposure condition as a between-subjects factor and distortion level and category as within-subjects factors. Although subjects in both conditions performed nearly perfectly on old patterns presented during the transfer test, those subjects trained only to criterion made more errors on old members (mean proportion = .058) than did the subjects who received an additional eight training trials (mean proportion = .018) [F(1,36) = 4.12, p < .05].

The error data on new transfer patterns are again pre-

sented in Table 1. The most important finding was that transfer errors decreased as a function of increased category exposure even though category size was held constant [$F(1,36) = 4.25, p < .05$]. In addition, transfer errors increased as a function of increasing distortion level [$F(3,108) = 67.50, p < .01$], and there was a significant category x distortion level interaction [$F(6,216) = 4.34, p < .01$]. The latter was due to: (1) performance on the prototype being worse for Category C than for Categories A and B, and (2) performance on medium distortions being better in Category B than in the other two categories.

GENERAL DISCUSSION

Previous studies varying category size have confounded this factor with category exposure. We obtained virtually no effect of category size when exposure was held constant, but we did obtain a clear effect of exposure when category size was held constant. These results call previous demonstrations of the category size effect into question. However, although its effect may be less pervasive than previously thought, it seems unlikely that category size never influences performance. For example, some within-category variability is needed to learn a concept, and variability normally should increase with increasing category size. It is possible that the effect of category exposure is stronger than that of category size primarily when there is appreciable variability within a category regardless of its size. Future research will have to separate the effects of category size and category exposure in order to discover when and how they each may affect the formation of ill-defined concepts.

With respect to theory, our findings pose a problem both for theories emphasizing the role of stored training stimuli in the classification of new stimuli and for prototype theories. The effects of category exposure suggest that performance is influenced by the amount of experience with the category rather than by the amount of experience with each individual training stimulus. Thus, the effect of exposure in our experiments seems to have depended on category-level information rather than on information about individual exemplars, a conclusion that seems inconsistent with theories that attribute classification of new transfer stimuli to their similarity to specific stored training exemplars. However, prototype theories also seem to need further development if they are to provide an adequate account of exposure effects. Currently, the latter theories provide little reason to expect that transfer performance would be facilitated appreciably by overtraining on a restricted set of stimuli.

The need for additional theoretical development in order to account for the effect of an acquisition variable like category exposure seems to be a natural outcome of the currently popular approach to studying ill-defined concepts. The central concern has been the question of what people learn. Theoretical developments have been most explicit regarding the retrieval processes and decision rules that subjects use in classifying new stimuli on the basis of whatever they may have learned (e.g., Hayes-Roth & Hayes-Roth, 1977; Medin & Schaffer, 1978; Reed,

1972). Research has emphasized the relationship between training stimuli and test stimuli (e.g., Chumbley, Sala, & Bourne, 1978; Goldman & Homa, 1977; Kellogg, 1981), as well as variables such as category size that influence category structure during training (e.g., Barresi, Robbins, & Shain, 1975; Homa et al., 1973; Homa & Vosburgh, 1976; Peterson, Meagher, Chait, & Gillie, 1973). Relatively little attention has been paid to acquisition processes. However, like retrieval and decision processes, acquisition processes may have important implications for the nature of what people learn when they acquire ill-defined concepts.

REFERENCES

BARRESI, J., ROBBINS, D., & SHAIN, K. (1975). Role of distinctive features in the abstraction of related concepts. *Journal of Experimental Psychology: Human Learning and Memory*, **104**, 360-368.

CHUMBLEY, J. J., SALA, L. S., & BOURNE, L. E. (1978). Bases of acceptability ratings in quasinatural concept task. *Memory & Cognition*, **6**, 217-226.

GOLDMAN, D., & HOMA, D. (1977). Integrative and metric properties of abstracted information as a function of category discriminability, instance variability and experience. *Journal of Experimental Psychology: Human Learning and Memory*, **3**, 375-385.

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.

HOMA, D., & CHAMBLISS, D. (1975). The relative contributions of common and distinctive information on the abstraction from ill-defined categories. *Journal of Experimental Psychology: Human Learning and Memory*, **104**, 351-359.

HOMA, D., CROSS, J., CORNELL, D., GOLDMAN, D., & SHWARTZ, S. (1973). Prototype abstraction and classification of new instances as a function of the number of instances defining the prototype. *Journal of Experimental Psychology*, **101**, 116-122.

HOMA, D., & VOSBURGH, R. (1976). Category breadth and the abstraction of prototypical information. *Journal of Experimental Psychology: Human Learning and Memory*, **3**, 322-330.

KELLOGG, R. T. (1981). Feature frequency in concept learning: What is learned? *Memory & Cognition*, **9**, 157-163.

MEDIN, D. L., & SCHAFFER, M. M. (1978). Context theory of classification learning. *Psychological Review*, **85**, 207-238.

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

PETERSON, M. J., MEAGHER, R. B., CHAIT, H., & GILLIE, S. (1973). The abstraction and generalization of dot patterns. *Cognitive Psychology*, **4**, 378-398.

POSNER, M. I. (1969). Abstraction and the process of recognition. In G. H. Bower & J. T. Spence (Eds.), *Psychology of learning and motivation* (Vol. 3). New York: Academic Press.

POSNER, M. I., GOLDSMITH, R., & WELTON, K. E. (1967). Perceived distance and the classification of distorted patterns. *Journal of Experimental Psychology*, **73**, 28-38.

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