Title
Effects of Category-Learning on Categorization An Analysis of Inference-Based and Classification-Based Learning

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 17(0)

Authors
Yamauchi, Takashi
Markman, Arthur B.

Publication Date
1995

Peer reviewed

Effects of Category-Learning on Categorization — An Analysis of Inference-Based and Classification-Based Learning

Takashi Yamauchi
Psychology Department
Columbia University
New York, NY 10027
takashi@psych.columbia.edu

Arthur B. Markman
Psychology Department
Columbia University
New York, NY 10027
markman@psych.columbia.edu

Abstract

It is widely acknowledged that categories have many functions, but few studies have actually addressed the impact of these functions on the way categories are learned. For instance, many categorization experiments predominantly rely on classification-based incremental learning. The problem with this approach is that it implicitly assumes that the function of categorization is separable from the way that categories are learned. In this study, we examined the relation between learning and the subsequent use of categories by contrasting three types of category-learning methods — inference-based, classification-based, and a combination of these methods. The results of the experiment indicate that there is an intricate relationship between category-learning and subsequent use of the category. The results further suggest that different processing modes may have been adopted by subjects in the different learning conditions.

The use of categories in a natural setting encompasses diverse functions including communication, analysis, comprehension, inference and comparison. An art historian, for instance, classifies paintings based on their artistic styles, and forms categories such as Gothic, Baroque, Rococo and Romanticism in order to analyze the difference between them. Categories may be formed for inference as well. A person may be classified as a Democrat or a Republican by considering her stand on abortion: Similarly, her political affiliation may be determined by reflecting on her personality. As these examples illustrate, categories serve various cognitive functions.

Despite the multifaceted use of categories, classification has been the dominant procedure employed to study categories (though, see Estes (1994) for some alternatives). A typical categorization experiment consists of two parts, a learning and a transfer phase. During learning, subjects are taught to classify stimuli into groups. Following learning, other measures of category structure like typicality ratings are used. This regular paradigm is based on the two assumptions: (1) the characteristics of category acquisition can be reliably measured by examining subjects' behavior in a classification task; (2) categorization processes are independent of the context in which the categories are learned. The potential danger in this research strategy is that category learning, when reduced to the acquisition of category names, may bear little resemblance to the acquisition of natural categories. While the representation of categories acquired through classification learning alone may be rich enough to support multiple functions of category use, it is also possible that the manner in which we use a category affects what is learned about it. In other words, the type of processing and representation used for category functions may be partly determined by the characteristics of learning.

There is some work that is consistent with the hypothesized link between the way categories are learned and what is acquired. Medin and Smith (1981) showed that different strategies adopted during a learning period resulted in qualitatively different performance in a transfer task. Elio and Anderson (1984) demonstrated that characteristics of stimulus presentation interacted with analytic and non-analytic modes of category processing. If the nature of categorization is determined by the characteristics of category learning, the mechanism of categorization cannot be adequately explained without analyzing the relation between category learning and the subsequent representation of categories.

The purpose of this study is to investigate the relationship between category learning and subsequent mechanisms of category use by introducing a new inference-based learning method. Among the many functions of categories, inference plays a central role (Anderson, 1990; Heit 1993; Holland, Holyoak, Nisbett & Thagard, 1986; Murphy & Ross, 1994; Rosch, 1976). Holland et al. (1986) point out that categories function to give goal-relevant expectations about instances, and Murphy and Medin (1985) suggest that categories are obtained as a consequence of inference, implying an alternative method to study categorization. We compared three types of learning procedures — standard classification-based learning, inference-based learning, and a mixture of classification and inference learning. In so doing, we examined the effects produced by inference and classification learning on the system of categorization.

The learning phase of this study consists of one of three learning conditions — Inference-only, Classification-only, or Inference-and-Classification conditions. The Classification-only condition is equivalent to a standard

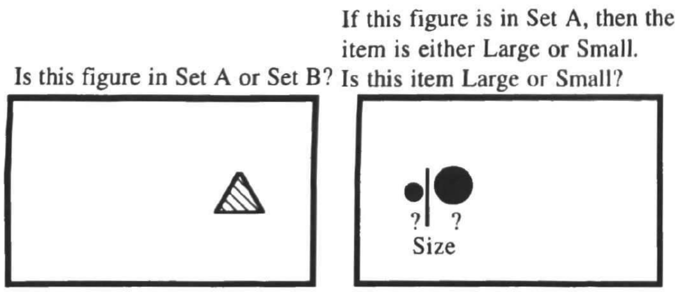

Figure 1: (a) A sample stimulus used in the experiment. (b) A sample stimulus used in a Critical-Feature inference. The choice of a subject reflects either the prototype stimulus (A0 — large circle) or the exemplar (A3 — small circle).

category-learning format; subjects view a stimulus and classify it with feedback given after each trial. In the Inference-only condition, subjects are instructed to infer one of two feature values of a stimulus along some dimension given its category label and other feature information. For instance in Figure 1b, the stimulus is drawn with a specific form, position and color, but the size of the figure must be inferred. The Inference-and-Classification condition is a mixture of the two described above.

Experiment

Method

Participants. Seventy-two subjects participated in this study (24 in each condition). They were recruited from the Columbia University community and were paid $6.00 for their participation.

Materials. The stimuli used for this experiment were like those used in the first experiment of Medin and Schaffer's (1978) studies. Specifically, the stimuli were geometric figures having four feature dimensions — form, color, size and position (Figure 1). The figure was placed on the left or the right side of a 20.3 x 17.4 cm rectangular frame drawn with a solid black line on the computer screen.

The structure of the two categories is illustrated in Table 1 (see Medin, Wattenmaker & Hampson, 1987). Eight stimuli (A1-A4, B1-B4) were divided into two categories. As the figure shows, each category is predominant in one of the two values. None of the stimuli, however, perfectly matches the dominant value of each category, so that no single feature can unambiguously determine the category division. The stimuli A0 and B0, which share the most common values with their category members, and the least common values with the contrasting category, can be considered prototypes of each category. These stimuli appeared only in the transfer tasks.

Procedure. The basic procedure of the experiment involved three phases — an initial learning phase, a distractor task and a final transfer phase.

Table 1. Category structure used for the experiment. F, S, C, and P stand for the four feature dimensions — form, size, color and position respectively, each of which has binary feature values — (1, 0) = F(circle, triangle), S(large, small), C(red, green) and P(left, right). The stimuli, A1-A4 and B1-B4, were shown in the learning task. All the stimuli, A0-A4 and B0-B4 were shown in the transfer phase. The values typed with the bold format were called "Critical-features."

	Set A					Set B			
	F	S	C	P		F	S	C	P
A1	1	1	1	**0**	B1	0	0	0	**1**
A2	1	1	**0**	1	B2	0	0	**1**	0
A3	1	**0**	1	1	B3	0	**1**	0	0
A4	**0**	1	1	1	B4	**1**	0	0	0
A0	1	1	1	1	B0	0	0	0	0

In the initial learning phase, subjects were randomly assigned to one of three experimental conditions — Classification-only, Inference-only and Inference-and-Classification. For all the three conditions, the entire learning phase consisted of 18 blocks with 8 trials in each block. Each stimulus appeared once in each block. The order of stimulus presentation was determined randomly.

In the Classification-only condition, subjects were shown one of the eight stimuli and were asked to indicate the category to which it belonged by clicking a button with the mouse. Initially, no information about the category division was given to subjects; and so subjects had to guess. Following subjects' response, feedback was provided; the stimulus and the feedback remained on the screen for three seconds after their response.

Subjects in the Inference-only condition were asked to infer the value of one of the four features of an item given its category label and information about the other three features. The stimuli in this condition were depicted with four features — form, color, size and position. However, on each trial the value of one feature was unspecified. For instance, on a given trial the color of the stimulus might be missing while the other three features and the category label were completely specified. On another trial, the size of one stimulus could be left unspecified while all the other information was provided (Figure 1b). At the onset of a trial, a stimulus was displayed on the screen, and subjects chose one of two values of the missing feature by clicking one of two labeled buttons with the mouse. In this study, subjects answered all the feature questions related to each stimulus except for the questions associated with the "exception value" (e.g. the feature value 0 in the category containing predominantly 1s).[1] For instance, in the stimulus A1, the feature inferences about form, color, and

[1] These questions were removed from the learning phase to be used for the Critical-feature transfer task. The detail of this transfer task (Critical-feature inference task) is described at the end of Procedure section.

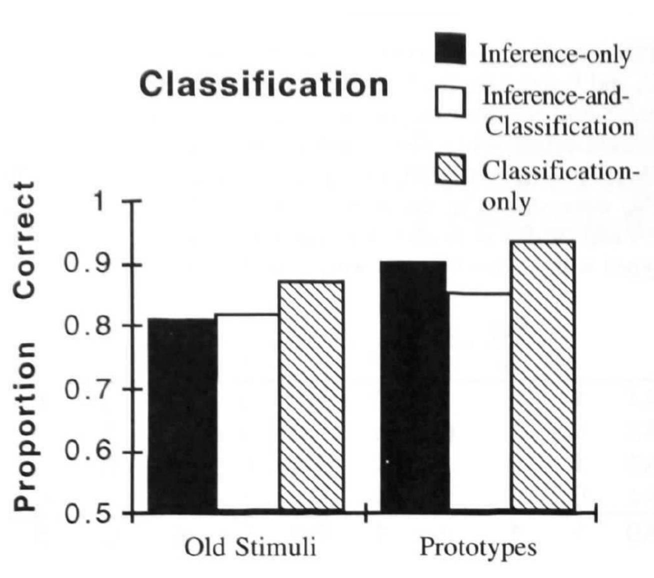

Figure 2: The performance for the classification.

size were asked but the inference about position was omitted.

The Inference-and-Classification condition was a mixture of classification and inference blocks. Half of the learning blocks were inference trials, and the remaining were classification. The order of blocks was determined randomly.

Following the learning trials, all the subjects participated in a distractor task in which they judged the pronounceability of nonsense words. This phase lasted about ten minutes.

After the distractor phase, subjects participated in the transfer tasks, which were the same for all groups. In this phase, both classification and inference tasks were given. Subjects were asked to make their decisions based on the categories learned during the initial phase. First, subjects were shown ten stimuli in sequence: the eight old stimuli that appeared in the learning phase and two new prototype stimuli (A0 and B0). They classified each stimulus without feedback. They also indicated whether they had seen the stimulus during the learning trials.[2] Following the classification test, subjects proceeded to the inference task in which they inferred the value of one of four features of stimulus given its category label and other three features. They performed all possible feature inferences. No feedback was given during the transfer. The entire experiment took 30 to 40 minutes.

Four dependent measures served for our analyses — learning rate of three learning conditions, classification and inference performance, and Critical-feature inference. Learning rate of three learning conditions was measured by calculating the average performance in the last three blocks in the learning phase. The performance of classification and inference was obtained from the transfer task in which all the subjects were exposed to the identical tasks. Of the four dependent variables, the special dependent measure, the Critical-feature inference, needs a further explanation. The Critical-features, which are typed with the bold format in Table 1, constitute the irregular values of each category. For instance, the values of all the features in Set A are 1 except the values of the Critical-features which are 0. In the Critical-feature inference trials, subjects made inferences to these "exception features". For instance, in stimulus A1 subjects were presented the information about form, size and color but the value of position was left for subjects to infer. There are two possible choices in A1, 1 or 0 (left or right). A choice of 0 would match the actual exemplar (A1), but a choice of 1 would be congruous with the prototype of the category (A0). In other words, inferences based on the exemplar learned would lead subjects to choose 0, but inferences based on the prototype of the category would lead subjects to choose 1. Consequently, the choice of feature values on these trials may indicate how the inferences were made.

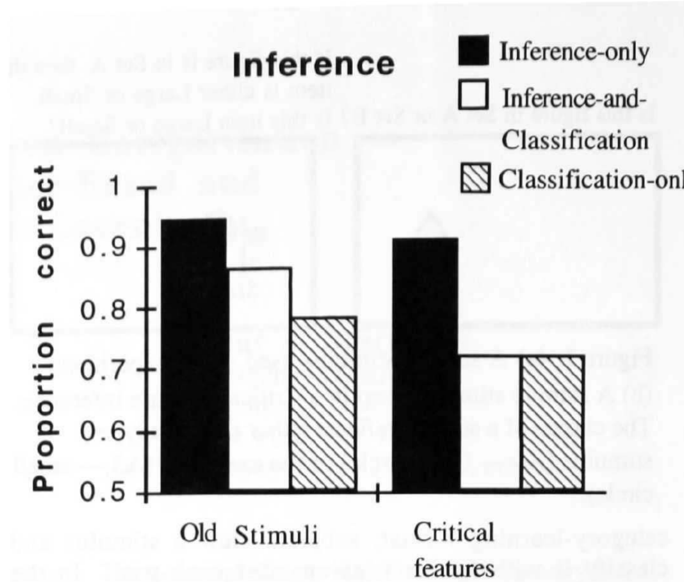

Figure 3: The performance for the feature inference. The ordinate of the Critical-feature measure stands for the proportion of prototype-accordance feature inference.

Results

The central results of this experiment are illustrated in Figures 2 and 3. All dependent measures were analyzed with one-way ANOVAs. In order to ensure that the pattern of the observed responses was not a simple reflection of the learning level associated with each condition, we examined the data separately for all the subjects who averaged above 90% correct responses in the last three blocks of the learning trials. First, we analyzed the learning rate for each condition by calculating the average performance in the last three blocks in the learning phase. The data showed that the subjects in the Inference-only condition obtained a higher

---

[2] We collected the recognition performance data of the subjects on an exploratory basis. Because this experiment was not designed to survey recognition performance (i.e., there were only two new stimuli out of ten stimuli), the data did not yield any meaningful results for further analyses.

learning rate (m= 93.3%) than did the subjects in either the Inference-and-Classification condition (m= 80.2%) or the Classification-only condition (m= 85.4%); $F(2, 69) = 2.795$, $p < 0.07$. Planned comparisons after the Bonferonni adjustment indicated that the difference between the Inference-only and the Inference-and-Classification condition was marginally significant; $t(46) = 2.01$, $p < 0.08$. No other comparisons approached significance. Overall, 19 subjects in the Inference-only condition averaged above 90% in the last three blocks of the learning phase, 10 subjects reached this criterion in the Inference-and-Classification condition, and 15 subjects reached this criterion in the Classification-only condition.

The classification transfer results are shown in Figure 2. Unsurprisingly, the data revealed that subjects in the Classification-only condition classified the old stimuli more accurately (m=87.5%) than did the subjects in the Inference-and-Classification (m=81.8%) and in the Inference-only condition (m=81.3%). Interestingly this difference was not significant; $F(2, 69) < 1$. The same measure of the performance in the subjects above 90% of learning level showed a similar trend: Classification-only (m= 95%), Inference-and-Classification (m= 93.8%) and Inference-only (m= 86.8%). As shown in Figure 2, all three groups were quite accurate in classifying prototype stimuli.

The inference performance for old stimuli indicated that the subjects in the Inference-only condition predicted feature values of old stimuli more accurately (m= 94.4%) than did the subjects in the Inference-and-Classification (m= 86.3%) or in the Classification-only condition (m= 78.5%); $F(2, 69) = 4.54$, $p< 0.02$. Planned comparisons revealed a significant difference between the Inference-only and the Classification-only condition; $t(46) = 2.68$, $p< 0.05$. No other differences reached significance. The performance of the subjects above a 90% learning level showed a similar pattern; Inference-only (m= 99.3%), Inference-and-Classification (m= 97.5%) and Classification-only (m= 87.2%). The differences observed in the three learning conditions clearly reflect the impact of particular types of category-learning on the transfer task.

The results obtained in the Critical-feature inference task suggested that inferences were made in accordance with the prototype more often than would be expected by chance for all three conditions. However, the subjects in the Inference-only condition tended to select feature values in accordance with the prototypes (m=91.1%) more often than did the subjects in the Inference-and-Classification (m=71.9%) or in the Inference-only condition (m= 71.9%); $F(1, 69) = 4.36$, $p < 0.02$. Planned comparisons after the Bonferroni adjustment showed that the difference between the Inference-only and the Classification-only condition as well as the difference between the Inference-only and the Inference-and-Classification condition were significant, respectively; $t(46) = 2.295$, $p < 0.05$ (Figure 3).[3] The performance of the subjects above a 90% learning level showed a similar trend; Inference-only (m= 94.7%), Inference-and-Classification (m= 80%) and Classification-only (m= 0.75.8%).

General Discussion

This experiment suggests that there is an intricate relationship between category learning and subsequent transfer tasks. First, all learning conditions provided subjects with some ability to use categories flexibly beyond the manner in which they were learned. Subjects in the Classification-only condition could still perform the inference task at an above-chance level, and subjects in the inference condition could perform the classification task at an above-chance level. These results reveal that both classification and inference learning lead to a representation flexible enough to cope with different cognitive tasks.

Despite the fact that all the learning conditions led to above-chance levels performance in both classification and inference tasks, performance in transfer tasks was generally better when the learning task matched the transfer task. It may be the case that each learning condition induced subjects to adopt different processing modes in the transfer tasks. The results of this experiment were not consistent with the view that the category label is just another feature of the category and that classifying an exemplar is equivalent to inferring the label given the other features (e.g., Anderson, 1990). In particular, the mixed inference and classification learning task was much more difficult than the inference based learning task. If classification involved just another inference, then these tasks should not reveal so great a difference in difficulty.

The results of Critical-feature inference trials provide an opportunity to examine the kinds of representations used as the basis of inferences. Strikingly, in all conditions, subjects' Critical-feature inferences were congruent with the prototypes, even though the actual exemplars seen suggested the opposite response. The level of prototype responding was highest in the conditions involving inference during learning. Whether this performance reflects a bias toward prototype formation in learning tasks involving inference, or simply better performance by subjects who did inference trials during learning is an open research question.

This study raises a number of interesting questions for further research. Categorization studies so far have been dominated by a classification-based learning format. As shown in this experiment, the performance in the transfer tasks mirrored the method adopted in learning. Thus when categories are learned and examined by classification tasks alone, the obtained results may embody characteristics

[3] The difference between the Inference-only and the Inference-and-Classification condition cannot be solely attributable to the overall difference of the inference performance between the two. Subjects in the Inference-and-Classification condition inferred the feature values of old stimuli 86.3% of the time. This number is significantly higher than their performance in the Critical-feature inference, (m= 71.9%); $t(46) = 2.21$, $p < 0.02$.

specific to classification but not to categorization. Prior studies have analyzed category functions primarily based on classification learning. These findings may be more relevant to classification processes than to categorization in general. We suggest that categorization mechanisms such as exemplar-based or prototype-based models need to be reexamined with respect to learning procedures.

Finally, although this study highlights the complex relationship that exists between category learning and category acquisition, it is far from clear precisely how category learning is related to the mechanisms of category acquisition. We expected that the Inference-and-Classification learning condition would yield good performance in both the transfer tasks. Contrary to our prediction, the Inference-and-Classification condition resulted in moderate levels of performance as compared to other learning conditions. In addition, the subjects in the Inference-and-Classification condition had the lowest level of learning rate. These results may be due to conflicts in the task demands of classification and inference learning. Classification, which guides subjects to acquire category labels pertaining to each stimulus, may lead subjects to seek a simple link between an instance and its label. Inference, on the other hand, may require encoding relations and correlations between features. Conflicts in processing demands may result in poor performance by subjects during both inference-based and classification-based learning. This question needs to be addressed in future studies in order to illuminate the impact of a multifaceted set of learning tasks on category formation.

Acknowledgments

We wish to acknowledge the support of Dendre Gentner, Trisha Lindemann, Valerie Makin, Douglas Medin, and Yung-Cheng Shen, who gave us valuable comments and suggestions.

References

Anderson, J. R. (1990). *The Adaptive Character of Thought* Lawrence Erlbaum Associates, Hillsdale, New Jersey 1990.

Elio, R., & Anderson, J. R. (1984). The effects of information order and learning mode on schema abstraction. *Memory & Cognition*, 12, 20-30.

Estes, W. K. (1994). Classification and Cognition. Oxford University Press, New York 1994.

Heit, E. (1992). Categorization using chains of examples. *Cognitive Psychology*, 24, 341-380.

Holland, J. J., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning and discovery.* Cambridge MA: MIT Press.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification. *Psychological Review*, 85 (3), 207-238.

Medin, D. L., & Smith, E. E. (1981). Strategies and classification learning. Journal of Experimental Psychology: Human Learning and Memory, 7, 241-253.

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

Murphy, G. L., & Medin, D. L. (1985). The roles of theories in conceptual coherence. *Psychological review.* Vol. 92. 289 316.

Murphy, G. L., & Ross, B. H. (1994). Predictions from uncertain categorizations, *Cognitive Psychology*, 27, 148-193.

Rosch, E., Mervis, C. B., Gray, W. B., Johnson, D. M., & Boyes-Braem, P. (1976). Basic Objects in Natural Categories. *Cognitive Psychology*, 8, 382-439.