Psychonomic Bulletin & Review
2002, 9 (4), 829-835

Comparing supervised and unsupervised category learning

BRADLEY C. LOVE
University of Texas, Austin, Texas

Two unsupervised learning modes (incidental and intentional unsupervised learning) and their relation to supervised classification learning are examined. The approach allows for direct comparisons of unsupervised learning data with the Shepard, Hovland, and Jenkins (1961) seminal studies in supervised classification learning. Unlike supervised classification learning, unsupervised learning (especially under incidental conditions) favors linear category structures over compact nonlinear category structures. Unsupervised learning is shown to be multifaceted in that performance varies with task conditions. In comparison with incidental unsupervised learning, intentional unsupervised learning is more rule like, but is no more accurate. The acquisition and application of knowledge is also more laborious under intentional unsupervised learning.

Categorization subserves many facets of cognition such as decision making, reasoning, object recognition, and aspects of language processing. Despite the intimate connections between categorization and many other aspects of cognition, the study of category learning has focused on one supervised learning task—namely, classification learning (Nosofsky, Gluck, Palmeri, McKinley, & Glauthier, 1994; Shepard, Hovland, & Jenkins, 1961). In supervised classification learning, the learner is given a stimulus, classifies it, and is provided with corrective feedback. Certainly, there are many other ways to learn about categories. For instance, some category learning tasks do not involve supervision. Supervised classification learning stresses one set of component processes out of many possible sets. Equating category learning with supervised classification learning (or any one induction task) would result in a limited understanding of category learning that does not extend across a wide range of situations (Love, 2001; Schank, Collins, & Hunter, 1986).

Work in supervised learning that has compared different learning modes has found strong interactions between different learning problems and learning modes, even when the learning modes are informationally equivalent (Goldstone, 1996; Yamauchi, Love, & Markman, 2002; Yamauchi & Markman, 1998). In other words, different supervised induction tasks stress different aspects of the stimulus set and perhaps invoke different processes. The aforementioned work points toward the multifaceted nature of supervised learning and, coupled with modeling work that clarifies the conceptual relations between the different learning modes (Love, Markman, & Yamauchi, 2000), provides for a more general understanding of category learning.

Instead of focusing on supervised learning, the work presented here examines two unsupervised learning modes (incidental and intentional unsupervised learning) and their relation to supervised classification learning. Depending on the task and the learner's goals, a learner can spontaneously develop categories (unsupervised learning) or conceptual organization can be strongly constrained by feedback (supervised learning). Developing a better understanding of unsupervised learning and its relation to the large body of work in supervised classification learning is critical given that a great deal of human learning may be unsupervised.

Learning Modes and Category Structures: Predictions

Supervised classification learning is intentional and encourages subjects to search for rules and perform hypothesis testing. Subsequently, learners exhibit sharp drops in their error rates (all-or-none learning) and are consciously aware of the rules[1] that they are entertaining (see Nosofsky, Palmeri, & McKinley, 1994). Accordingly, supervised classification problems are learned most efficiently when the category structure can be described by a simple rule. The results from Shepard et al.'s (1961) six[2] problems follow this pattern (see Table 1 for problem descriptions). The Type I problem (in which a rule can be formed on the first stimulus dimensions) is learned faster than the Type II problem (in which a rule can be formed on the first two stimulus dimensions), which is learned faster than the Type IV problem (in which the dimensions are intercorrelated, but the learner must attend to all stimulus dimensions), which is learned faster than the Type VI problem (in which no exploitable regularity exists and the stimuli must be memorized).

This work was supported by AFOSR Grant F49620-01-1-0295 to B.C.L. I thank Todd Maddox, Art Markman, Doug Medin, and Yasu Sakamoto for helpful discussions. Correspondence concerning this article should be addressed to B. C. Love, Department of Psychology, University of Texas, Austin, TX 78712 (e-mail: love@psy.utexas.edu or http://love.psy.utexas.edu/).

Table 1
The Logical Structures of the Type I, II, IV, and VI Learning Problems Considered by Shepard, Hovland, and Jenkins (1961)

Item	I	II	IV	VI
1 1 1	1	1	1	1
1 1 2	1	1	1	2
1 2 1	1	2	1	2
1 2 2	1	2	2	1
2 1 1	2	2	1	2
2 1 2	2	2	2	1
2 2 1	2	1	2	1
2 2 2	2	1	2	2

Note—Each learning problem involves assigning eight stimulus items to one of two categories. Each item varies on four binary valued stimulus dimensions (e.g., size: small or large). The displayed value of these dimensions is denoted by a 1 or a 2. The relation of the fourth dimension (i.e., the category) to the other three dimensions defines the category structure of each learning problem.

These findings have been very influential in developing theories and models of supervised classification learning (Nosofsky, Gluck, et al., 1994). What is curious about these findings is that the Type II problem, which has a highly nonlinear category structure, is learned faster than the Type IV problem, which has a family resemblance structure (a linear category structure in which all dimensions are intercorrelated and each dimension provides imperfect evidence for category membership). Natural categories are thought to have a family resemblance structure (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976; Wittgenstein, 1953), yet such categories are not readily acquired through supervised classification learning.

One possible explanation for this discrepancy is that supervised classification learning is not typical of real-world learning situations. After all, the world does not always respond to the queries of the learner. In many situations, learning is not even one's primary objective when one interacts with a stimulus; instead, it is a consequence of completing another task (i.e., learning can be incidental). It is quite possible that supervised classification learning and unsupervised learning (particularly incidental unsupervised learning) lead to very different patterns of acquisition.

According to Wattenmaker (1991), it can be predicted that intentional learning (both supervised and unsupervised) will promote rule formation, whereas incidental learning will promote similarity-based processing (e.g., similarity to stored exemplars). These two basic processes loosely correspond to Sloman's (1996) proposal for dual conceptual systems (rule vs. similarity based). Although the linear/nonlinear distinction has not proved very meaningful in supervised classification learning (Medin & Schwanenflugel, 1981), this distinction could prove crucial in unsupervised learning. It is predicted that subjects learning through incidental unsupervised learning will show a preference for family resemblance structures (e.g., the Type IV problem) and will show deficits for nonlinear category structures (e.g., the Type II problem). Intentional unsupervised learning should more closely match supervised classification learning performance. Along these lines, it is expected that learners will engage in more rule construction in intentional than in incidental unsupervised learning. Rule application is generally regarded as slower than similarity-based retrieval process (see Sloman, 1996, for a review), therefore test performance should be slower under intentional unsupervised learning (although the complexity of the stimuli and the rules could affect the relative speeds of these processes). Another consequence of rule application is increased variability in performance across subjects. Subjects who construct successful rules will perform at high levels, whereas subjects who construct rules that are orthogonal to the actual category structure will perform at low levels. Therefore, greater variability should be observed under intentional learning conditions.

Overview of Experiments 1 and 2

One unique feature of the present experiments is that they allow for direct comparisons of unsupervised and supervised learning performance. In order to directly compare unsupervised and supervised learning, a comparable dependent measure is needed. Ideally, the dependent measure would closely correspond to training accuracy, which is a measure widely used in assessing the difficulty level of supervised classification learning problems (e.g., Nosofsky, Gluck et al., 1994; Shepard et al., 1961).

In order to develop such a measure, stimuli were created by embedding the category label (typically, subjects classify each stimulus as a member of Category A or B) into each stimulus as a fourth binary-valued perceptual dimension (see Table 1). On supervised classification study phase trials, the subjects were shown the value of the first three perceptual dimensions and were queried on the fourth (e.g., 1 1 2 ?). After the subjects responded, feedback was provided and the complete stimulus description was shown (in the case of the Type VI problem, 1 1 2 2 was shown). In unsupervised learning, all four perceptual dimensions were shown on study phase trials (the fourth dimension was not queried).

In intentional unsupervised learning, the subjects were aware that they were involved in a learning task, and their efforts were directed toward the learning task (i.e., they actively searched for patterns that characterized the training items). In incidental unsupervised learning, the subjects were not aware that they were in a learning task and they directed their effort toward another task (in this case, they rated how pleasant they found each stimulus). This incidental task is evaluative in nature and probably encourages subjects to encode all four stimulus dimensions. Other incidental tasks (evaluative or nonevaluative) could stress different aspects of the stimulus set and lead to disparate patterns of performance (see Love, in press, for a comparison of different incidental tasks). For example, if the incidental task was to indicate whether a stimulus was small or large, perhaps little information concerning the other stimulus dimensions and the relations across dimensions would be encoded.

Category learning performance was measured in the test phase (following the study phase). The test phase was identical for all conditions. The subjects viewed a pair of stimuli that varied only on the fourth dimension (i.e., the category dimension). The subjects were instructed to choose the item that appeared during the study phase. As in traditional supervised classification learning studies, the subjects could base their judgments on their knowledge of the relationship between the category dimension and the other dimensions (e.g., rules, correlations, etc.) as well as on memorized exemplars.

This testing procedure is novel, but it is comparable to the procedure used by Billman and Knutson (1996). Subjects in their study chose which of two stimuli was more familiar (i.e., similar to the study-phase stimuli). One stimulus preserved a studied correlation, whereas the other stimulus violated the correlation. In the present case, the "correlation" being tested was between the first three stimulus dimensions and the fourth. More broadly, the present testing procedure bears some resemblance to old/new recognition judgments, which are typically construed as measuring explicit memory (e.g., Knowlton & Squire, 1994), although recognition memory judgments also can be influenced by nonconscious fluency of processing (Johnston, Hawley, & Elliot, 1991; Seger, 1994). The present procedure differs from typical recognition judgments in some key ways. A forced choice procedure in which the two stimuli always differed in one respect (the category dimension) was used. In essence, on each test, the subject was asked to determine which category corresponded to the other three stimulus dimensions.

In these experiments, it is probably best not to view the fourth stimulus dimension as a category label per se. During the study phase of the unsupervised learning conditions, the subjects were not aware that the fourth stimulus dimension would be queried in the test phase. The study procedure placed no emphasis on the fourth stimulus dimension. In the unsupervised learning conditions, the test phase simply measured the extent to which the subjects grasped the relation between the first three dimensions and the fourth stimulus dimension (some relations or category structures may be more readily captured than others). Importantly, the subjects were free to organize the stimuli in any manner that they saw fit. For instance, the subjects could organize the stimuli along one dimension, as they typically do in sorting tasks (Medin, Wattenmaker, & Hampson, 1987). Such an organization would lead to chance performance in the test phase. In contrast to unsupervised learning, the supervised classification learning study phase is more constrained and requires subjects to predict the fourth stimulus dimension on the basis of the first three stimulus dimensions— a learning task that parallels the test phase task. Like the unsupervised learning modes, different relations between these first three dimensions and the fourth should affect test performance.

One important methodological claim is that test-phase accuracy is analogous to accuracy rate in supervised classification learning training, thus allowing for comparisons between unsupervised and supervised learning studies. Experiment 1 directly tested this claim in a supervised classification learning experiment. The claim is supported if study and test-phase accuracy are highly correlated.

EXPERIMENT 1

Method

Subjects. Two hundred fifty-two University of Texas undergraduate students participated for course credit. Each subject was randomly assigned to one of the four learning problems.

Apparatus. The experiments were run on Pentium III computers operating in DOS. The data were collected using an in-house real-time data collection system. The monitors had 15-in. CRT color displays and a refresh rate of 16.67 msec.

Stimuli. The study phase stimuli were geometric figures that varied in border color (yellow or white) and three of the four following binary-valued dimensions: size (small or large), color (blue or purple), texture (smooth or dotted), and diagonal cross (present or absent). Three of these four dimensions were randomly selected for each subject, with the remaining dimension fixed to one of its two values (also randomly determined). The three dimensions were mapped (randomly assigned to each subject) onto the logical structure shown in Table 1 with the border always serving as the fourth binary dimension (i.e., the category label). The assignment of dimension values was also random for each subject (e.g., for some subjects the value 2 on the size dimensions signified a large figure, for others it signified a small figure). The five stimulus dimensions were all equally salient and independent (as verified by multidimensional scaling of pairwise similarity ratings gathered from a previous study; see http://love.psy.utexas.edu/stimuli for details and to download the stimuli).

Procedure. In the study phase, the subjects were instructed to learn to predict the border color from the values of the other three stimulus dimensions. On each trial, the stimulus was presented with the border absent. The subjects pressed the "W" key to indicate that they believed the border was white and the "Y" key to indicate that they believed the border was yellow. After each response, the correct border was displayed (i.e., the entire stimulus was shown), and a positive tone sounded if the subject was correct, whereas a negative tone sounded if the subject was incorrect. The complete stimulus was displayed for 1,500 msec. A blank screen (i.e., a black screen) was then displayed for 1,000 msec and the next trial began. Unbeknownst to the subjects, response times were collected. Trials were organized into blocks (a block is the random presentation of each stimulus in a random order). The subjects completed 10 study-phase blocks.

After completing the study phase, the subjects were presented with a series of three arithmetic problems. Each problem consisted of two integers (randomly generated between 10 and 49) presented side by side (e.g., "22 + 34 = ?"). The subjects received both auditory and visual feedback indicating whether they had added the numbers correctly. The purpose of this filler task was to prevent the subjects from rehearsing information from the study phase.

After completing the three arithmetic problems, the subjects began the test phase. In the test phase, the subjects were presented with pairs of figures and had to choose which figure had been displayed in the study phase (forced choice old/new). One member of each pair was a stimulus that appeared in the study phase (the *old* figure), whereas the other member was a *new* figure that was identical to the old figure except that it displayed the opposite value on the fourth stimulus dimension (i.e., the category *label* dimension). On each test phase trial, the two figures were displayed side by side with the text "Old: left (Q) or right (P)?" displayed above the figures. The subjects pressed the "Q" key (on the left side of the keyboard) if they thought the figure on the left side of the screen was the old figure, whereas they pressed the "P" key (on the right side of the keyboard) if they thought the figure on the right was the old figure. After the subjects responded, a positive tone sounded and then a black screen with the string "Thank You" was displayed for 1,667 msec regardless of the correctness of their responses. A blank screen was then presented for 824 msec and then the next trial began. Whether the old item appeared on the left or right side of the screen was random for each trial. The subjects completed three test phase blocks (i.e., 24 trials).

Results and Discussion

The main results are illustrated in Table 2 under the heading Supervised Classification Learning. The predicted ordering (Type I, Type II, Type IV, Type VI) was observed in both the study and test accuracy data. All pairwise differences were significant at the .01 level, except for the differences between between the Type II and Type IV problems, which did not approach significance. Nevertheless, the data do reveal an advantage of Type II over Type IV.[3] For example, 21 out of 62 subjects in the Type II problem completed the test phase with greater than 95% accuracy compared with 5 out of 63 subjects in the Type IV problem [$\chi^2(1; N = 125) = 14.06, p < .001$].

The primary goal of Experiment 1 was to demonstrate that study-phase accuracy correlates with test-phase accuracy, thus enabling comparisons between supervised and unsupervised learning modes. The correlations were positive and significant. Averaging over subjects, the correlation between the mean study-phase accuracy and mean test-phase accuracy was .98 [$t(2) = 7.21, p < .05$]. The correlation over individuals (not groups) was .82 [$t(250) = 22.40, p \approx 0$].

EXPERIMENT 2

Method

Subjects. Three hundred eighty-five University of Texas undergraduate students participated for course credit. Each subject was randomly assigned to one of the eight conditions (one of the four learning problems under either incidental or intentional learning conditions).

Stimuli. The same stimulus set was utilized in Experiments 1 and 2. The one difference was that in Experiment 2, the category label dimension (the fourth stimulus dimension) was randomly selected for each subject (the supervised classification learning procedure utilized in Experiment 1 necessitated that the border color always be the fourth dimension).

Procedure. In the study phase of the incidental learning mode, the subjects were instructed to view a series of figures and rate the pleasantness of each figure on a 1–9 scale (1 was *unpleasant*; 9 was *pleasant*). The subjects were told that the purpose of the experiment was to norm the stimuli for a future experiment. The subjects completed 10 study-phase blocks. On each trial, the figure was shown on a black background for 1,000 msec. Then, the text *unpleasant* (1 to 9) *pleasant* was displayed above the figure. At this point, the subject could respond by pressing a key from 1 to 9. After the subject responded, a positive tone sounded and a blank (i.e., black) screen was displayed for 1,000 msec, followed by the next trial. Unbeknownst to the subjects, response times were collected.

In the study phase of the intentional learning mode, the subjects were instructed to memorize each figure. The subjects were told that searching for patterns across the figures might aid in memorizing the figures. The subjects completed 10 study blocks (the same number of trials as in the incidental conditions). On each trial, the figure was shown on a black background for 1,000 msec. Then, the text "Press the spacebar to continue" was displayed above the figure. At this point, the subject could respond by pressing the space bar. After the subject responded, a positive tone sounded and a blank (i.e., black) screen was displayed for 1,000 msec, followed by the next trial. Unbeknownst to the subjects, response times were collected.

The procedure for the arithmetic problems and the test phase was identical to that in Experiment 1. The three learning modes featured in the experiments differed only in the study phase.

Results and Discussion

The data of primary interest were the subjects' new/old forced choice judgments in the test phase. Figure 1 displays the mean accuracy level for the four problem types under both incidental and intentional learning conditions. The data are also shown in Table 2. Clearly, category structure (i.e., problem type) affected accuracy levels. Accuracy levels were significantly above the chance level of .5 with $p < .01$ for all conditions except for the Type VI problem under intentional learning conditions [$t(48) = 1.78, p \approx .08$]. When the Type VI problem is used as a baseline for item learning (as opposed to category learning)[4], the Type I and Type IV problems under incidental learning conditions and the Type I, II, and IV problems under intentional

Table 2
The Results From Experiments 1 and 2

	Study Accuracy		Study Time		Test Accuracy		Test Time	
Problem	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
			Incidental Unsupervised Learning					
Type I	NA		1,860	78	.85	.024	2,035	112
Type II	NA		1,691	72	.56	.018	2,250	153
Type IV	NA		1,799	73	.67	.016	2,233	154
Type VI	NA		1,899	136	.56	.017	2,112	134
			Intentional Unsupervised Learning					
Type I	NA		1,931	142	.84	.027	2,034	173
Type II	NA		2,333	224	.64	.024	2,640	254
Type IV	NA		2,355	179	.67	.018	2,641	160
Type VI	NA		2,593	292	.54	.020	2,433	175
			Supervised Classification Learning					
Type I	.86	.021	935	.64	.89	.025	1,636	103
Type II	.67	.021	1,678	.97	.73	.029	2,902	170
Type IV	.65	.013	1,446	.70	.70	.021	2,649	163
Type VI	.59	.015	1,693	.93	.61	.024	3,018	183

Note—Times are measured in milliseconds.

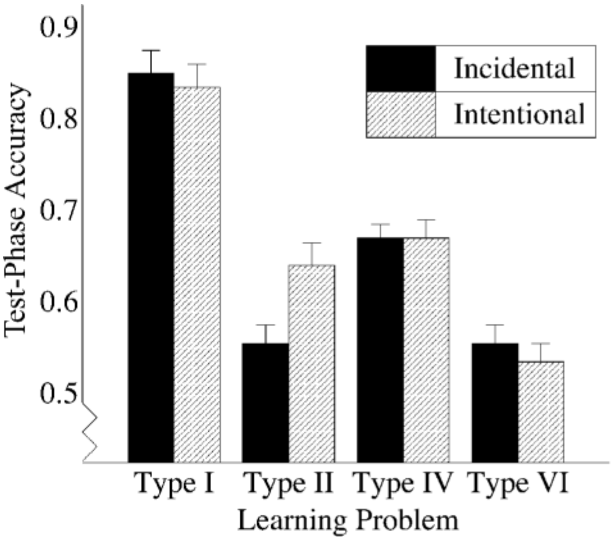

Figure 1. Mean test-phase accuracy and the standard error of the mean are displayed for the four problem types under both incidental and intentional learning conditions in Experiment 2.

learning conditions are all significantly different ($p <$ .001) than the Type VI problem under either incidental or intentional study conditions.

Comparing unsupervised learning to supervised classification learning. Both incidental and intentional unsupervised category learning led to patterns of acquisition that differ from supervised classification learning—neither unsupervised learning mode led to the Type I, Type II, Type IV, or Type VI difficulty ordering found in supervised classification learning. This discrepancy was greatest for incidental unsupervised learning—accuracy levels were substantially higher (.67 vs. .56) for the Type IV problem than for the Type II problem [$t(93) = 4.66, p \approx 0$].

Comparing incidental and intentional unsupervised learning. Important differences exist between incidental and intentional unsupervised learning. As predicted, the subjects performed at higher levels (.64 vs. 56) on the Type II problem (the low dimensional nonlinear problem) under intentional than under incidental learning conditions [$t(94) = 2.65, p < .01$]. This difference may reflect that incidental learning is better suited for linear category structures, whereas intentional learning is better suited for acquisition of categories that can be described by a compact rule (e.g., the Type II problem).

One predicted difference between incidental and intentional unsupervised learning is that intentional learning should promote explicit rule construction and rule application to a greater extent than would incidental learning. One manifestation of this prediction is that subjects should exhibit more all-or-none learning performance in the test phase under intentional conditions and this should lead to greater variability in subjects' performance. This prediction held. A paired comparison of the four learning problems revealed that the variance was larger (.025 vs. .018) in the intentional than in the incidental conditions [$t(3) = 4.66, p < .05$].

It was predicted that the greater prevalence of rule construction in the intentional conditions would also manifest itself in increased time per trial in the study phase, even though the intentional conditions only required space-bar presses during the study phase, whereas the incidental conditions required pleasantness ratings. This prediction held. The mean of each subjects' median response time during the study phase was 1,812 msec under incidental conditions compared with 2,303 msec under intentional conditions [$t(383) = 4.11, p \approx 0$]. The predicted prevalence of rule application in the test phase of the intentional conditions compared with the incidental conditions should have resulted in increased response times in the test phase as well. As predicted, the mean response time for the subjects under incidental conditions was 2,158 msec compared with 2,436 msec for intentional learning [$t(383) = 2.31, p < .05$].

Although not predicted, the subjects were more accurate in the arithmetic task (.92 vs. .87) after completing the incidental study phase than after completing the intentional study phase. These data were not normally distributed and were analyzed nonparametrically by classifying the subjects into two groups: those who made one or more errors and those who made no errors. In the incidental conditions, 149 out of 190 subjects were error free compared with 128 out of 195 subjects in the intentional conditions [$\chi^2(1; N = 385) = 7.17, p < .01$]. This result might reflect increased fatigue on the part of the subjects under intentional learning conditions or could indicate that these subjects attempted to rehearse the stimuli viewed in the study phase and that this interfered with arithmetic performance.

SUMMARY AND CONCLUSIONS

Recent work in supervised learning highlights the importance of comparing multiple learning modes in order to develop more general theories of learning (Love et al., 2000; Yamauchi et al., 2002). Experiments 1 and 2 extend this work by considering unsupervised learning modes. Acquisition patterns displayed by the subjects in supervised classification learning did not extend to unsupervised learning tasks. In particular, an advantage of linear category structures over nonlinear category structures was displayed in the unsupervised learning studies (particularly when the subjects were engaged in incidental unsupervised learning).

Important differences between incidental and intentional unsupervised learning were found. In only one case (the Type II problem) was there an advantage for intentional learning. This one advantage is compatible with the hypothesis that intentional learning encourages explicit rule formation, whereas incidental learning is better matched to the acquisition of linear category structures. This finding may partially explain why we have the natural categories that we have (assuming that incidental unsupervised learning is common in real-world learning situations).

Numerous disadvantages were observed for intentional unsupervised learning; presumably the drive to create and apply explicit rules led to increased study and test phase response times when compared with the incidental conditions. Interestingly, the subjects in the intentional conditions also performed worse in the filler task involving arithmetic problems, possibly indicating increased fatigue or the attempted rehearsal of study-phase items. In summary, there is no advantage to engaging in intentional unsupervised learning over incidental unsupervised learning except in cases in which the target concept is low dimensional and nonlinear (e.g., the Type II problem). All advantages for the intentional learning mode might disappear with higher dimensional stimuli because of the increased size of the hypothesis space. Overall, the results suggest that unsupervised learning is multifaceted and that it is incorrect to characterize unsupervised learning as stimulus driven, incremental, and passive (Berry & Dienes, 1993; Hayes & Broadbent, 1988; Hock, Malcus, & Hasher, 1986; Kellogg, 1982).

It might be tempting to map the incidental and intentional learning conditions onto work that posits multiple memory systems (e.g., Cohen & Eichenbaum, 1993; Squire, 1992). This may be a productive avenue to explore, but only if memory systems are defined in terms of the processes they subserve. Any successful memory systems account will have to include a role for how a stimulus is processed at encoding. There is ample evidence that the processes (or stimulus aspects) that are stressed at encoding play a major role in determining what is acquired (Markman & Makin, 1998; Roediger, Weldon, & Challis, 1989; Ross, 1996; Whittlesea & Dorken, 1993).

Shepard et al.'s (1961) work paved the way for quantitative modeling of supervised classification learning. Hopefully, the present work will lead to greater attention to the modeling of unsupervised learning and its relation to supervised learning modes such as supervised classification learning. The present data provide a valuable test bed for model development. Many existing models of supervised learning, such as GCM (Nosofsky, 1986), ALCOVE (Kruschke, 1992), the rational model (Anderson, 1991), and SUSTAIN (Love & Medin, 1998), can be applied to the present data with little or no modification. For example, the GCM can model the test data by calculating a measure of familiarity for test items by probing its exemplar memory of study items (see Nosofsky, 1988). On each test trial, the familiarity scores for the two alternatives can be fed into the GCM's choice function to determine the model's response. Modeling efforts along these lines are currently underway.

REFERENCES

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

Barsalou, L. W. (1990). On the indistinguishability of exemplar memory and abstraction in category representation. In T. K. Srull & R. S. Wyer (Eds.), *Content and process specificity in the effects of prior experiences: Advances in social cognition* (pp. 61-88). Hillsdale, NJ: Erlbaum.

Berry, D. C., & Dienes, Z. (1993). *Implicit learning: Theoretical and empirical issues*. Hillsdale, NJ: Erlbaum.

Billman, D., & Knutson, J. (1996). Unsupervised concept learning and value systematicity: A complex whole aids learning the parts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 458-475.

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, amnesia, and the hippocampal system*. Cambridge, MA: MIT Press.

Goldstone, R. L. (1996). Isolated and interrelated concepts. *Memory & Cognition*, **24**, 608-628.

Goldstone, R. L., & Kruschke, J. K. (1994). Are rules and instances subserved by separate systems? *Behavioral & Brain Sciences*, **17**, 405.

Hayes, N., & Broadbent, D. E. (1988). Two modes of learning for interactive tasks. *Cognition*, **28**, 249-276.

Hock, H. S., Malcus, L., & Hasher, L. (1986). Frequency discrimination: Assesing global and elemental letter units in memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 232-240.

Johnston, W. A., Hawley, K. J., & Elliot, J. M. (1991). Contribution of perceptual fluency to recognition judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 210-223.

Kellogg, R. T. (1982). When can we introspect accurately about mental processes? *Memory & Cognition*, **10**, 141-144.

Knowlton, B. J., & Squire, L. R. (1994). The information acquired during artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 79-91.

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

Love, B. C. (2001). Three deadly sins of category learning modelers. *Behavioral & Brain Sciences*, **24**, 687-688.

Love, B. C. (in press). The multifaceted nature of unsupervised category learning. *Psychonomic Bulletin & Review*.

Love, B. C., Markman, A. B., & Yamauchi, T. (2000). Modeling classification and inference learning. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence* (pp. 136-141). Cambridge, MA: MIT Press.

Love, B. C., & Medin, D. L. (1998). SUSTAIN: A model of human category learning. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence* (pp. 671-676). Cambridge, MA: MIT Press.

Markman, A. B., & Makin, V. S. (1998). Referential communication and category acquisition. *Journal of Experimental Psychology: General*, **127**, 331-254.

Medin, D. L., & Schwanenflugel, P. J. (1981). Linear separability in classification learning. *Journal of Experimental Psychology: Human Learning & Memory*, **7**, 355-368.

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

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

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 700-708.

Nosofsky, R. M., Gluck, M. A., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, **22**, 352-369.

Nosofsky, R. M., & Palmeri, T. J. (1996). Learning to classify integral-dimension stimuli. *Psychonomic Bulletin & Review*, **3**, 222-226.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, **101**, 53-79.

Roediger, H. L., III, Weldon, M. S., & Challis, B. H. (1989). Explaining dissociations between implicit and explicit measures of retention: A processing account. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 3-41). Hillsdale, NJ: Erlbaum.

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

Ross, B. H. (1996). Category representations and the effects of interacting with instances. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 1249-1265.

Schank, R. C., Collins, G. C., & Hunter, L. E. (1986). Transcending inductive category formation in learning. *Behavioral & Brain Sciences*, **9**, 639-686.

Seger, C. A. (1994). Implicit learning. *Psychological Bulletin*, **115**, 163-196.

Shepard, R. N., Hovland, C. L., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs*, **75** (13, Whole No. 517).

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, **119**, 3-22.

Squire, L. R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, **99**, 195-231.

Wattenmaker, W. D. (1991). Learning modes, feature correlations, and memory-based categorization. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 908-923.

Whittlesea, B. W. A., & Dorken, M. D. (1993). Incidentally, things in general are particularly determined: An episodic-processing account of implicit learning. *Journal of Experimental Psychology: General*, **23**, 227-248.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Oxford: Blackwell.

Yamauchi, T., Love, B. C., & Markman, A. B. (2002). Learning nonlinearly separable categories by inference and classification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 585-593.

Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory & Language*, **39**, 124-149.

NOTES

1. Whether the underlying representations that support rule-like performance are actually rules, as opposed to exemplars, prototypes, or clusters coupled with a selective attention mechanism, is not of primary interest here and may in some cases be indeterminable (Barsalou, 1990; Goldstone & Kruschke, 1994).

2. Only the Type I, II, IV, and VI problems are considered here. Performance on Types III and V is comparable with the Type IV problem.

3. The relatively small Type II advantage here (cf. Nosofsky, Gluck, et al., 1994) is due to the nature of the stimulus set and not to the present methodology. The same pattern of results is observed when the standard methodology (e.g., subjects predict membership in either Category A or B) is paired with the present stimulus set. However, typical stimulus sets (paired with the present methodology) do lead to a larger Type II advantage. Nosofsky and Palmeri (1996) provide another example of how Type II and IV performance can interact with the nature of the stimulus set. One difference between the present stimulus set and others is that the separation of dimension values (e.g., the distinction between a small stimulus and a large stimulus) is relatively small. The fact that such differences in the nature of the stimulus set can affect the difficulty level of learning problems has implications for theories of category learning. These implications are being systematically investigated. It should be stressed that the present stimulus set is normed and that each dimension is equally salient and independent. These steps should, but are usually not, taken by all researchers.

4. This distinction may not be meaningful on some views (e.g., an exemplar theory of category learning).

(Manuscript received August 23, 2001;
revision accepted for publication January 16, 2002.)