

# Specializing the Operation of an Explicit Rule

Scott W. Allen and Lee R. Brooks  
McMaster University, Hamilton, Ontario, Canada

The effect of practice on the operative form of a rule was investigated by giving subjects an easy, perfectly predictive classification rule, followed by training in applying that rule to a set of practice items. On a subsequent transfer test, the accuracy and speed of classifying new items was strongly affected by similarity to previously seen items, suggesting that the effect of practice was not simply to automatize the rule. The effect occurred with pictorial, easily integrated stimuli, but not with lists of verbally stated features. Subjects generally did not have insight into the role of previous items in their performance. This dependence on prior episodes may be frequent in ecologically common conditions and is of special interest when the categorization rule becomes uncertain, as when a rule has only heuristic value.

When dealing with a newly learned rule, one is often aware of relying on previous episodes of applying the rule. The specific materials or problems to which the rule was originally applied seem to have some privilege, particularly when the rule initially seems quite abstract. Understanding a difficult rule, or at least understanding how to apply it, seems to occur in terms of previous concrete applications. This impression is corroborated by Ross in his interesting work on reminders in problem solving (Ross, 1984, 1987, 1989; Ross & Kennedy, 1990). He demonstrated that performance while learning a word processing program or learning to solve simple probability problems is influenced by specific analogies with previously encountered problems.

But what happens to these prior episodes of problem solving when the rule is not difficult, or when one has had sufficient practice to make it no longer seem difficult? Introspectively, the prior examples seem to disappear from active processing, a suggestion reflected in a wide variety of theories in cognitive psychology. In this article we argue that under common conditions, prior episodes retain an important role in helping apply the rule to new material, well after clear awareness of the prior episodes has disappeared. We make an ecological argument for the conditions under which episode-based knowledge has continued value.

## Abstractive Generalization

An intuitively plausible view of the changes in processing that occur during practice with a rule is what we will refer to as *abstractive generalization*. With increasing practice, processing comes under the control of fast, general, and automatic procedures that concentrate solely on the relevant information given in the stimulus display. With practice, a rule that was explicitly known becomes automatized, consuming a decrease-

ing amount of processing resources, and with decreasing conscious access to the details (e.g., Shiffrin & Schneider, 1977; or the knowledge compilation process in ACT\* of Anderson, 1983). If the rule is not explicitly known, sufficient experience with categorized exemplars allows the essential structure of relevant variables to be learned and used implicitly (e.g., Broadbent, 1989; Broadbent, FitzGerald, & Broadbent, 1986; Lewicki, 1986; Reber, 1989). In the case of natural categories, sufficient experience with exemplars causes enough predictors to be learned to allow categorization by an additive (linearly separable) rule or by an extracted prototype. For example, it has been argued that with ill-defined categories, such as polygon stimuli (Homa, 1978; Homa, Sterling, & Treppel, 1981) and Impressionist paintings (Hartley & Homa, 1981), practice leads to a decline in the importance of previously encountered instances and an increase in the importance of abstracted information regarding central tendencies. In all of these cases, the changes with practice could be called analytic because they describe a process whereby the stimulus is decomposed into categorically relevant components.

Certainly abstractive generalization accords with the introspective observation that one is less aware of prior processing episodes after practice. Further, having an efficient procedure for applying a rule may require less storage capacity and less search time than would a large number of redundant subprocedures.

## Specialization of a Procedure

The opposite change—specialization—also occurs with practice. As a result of experience with particular common cases, a known procedure can be specialized into several forms, each of which is less general than the original. The most obvious case is one in which the original rule or procedure must be amended to produce an accurate answer for an exceptional case. But specialization also occurs in response to mere difficulty in application, such as, in probability, using separate formulae for the special case of statistical independence, rather than relying on the more complicated covariation formulae. The motive for the specialization in each of these cases is difficulty: difficulty that results in actual errors, diffi-

---

This research was carried out with the support of the Natural Sciences and Engineering Research Council of Canada, through a graduate scholarship to Scott W. Allen and a grant to Lee R. Brooks.

Correspondence concerning this article should be addressed to Lee R. Brooks, Department of Psychology, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4M8. Electronic mail may be sent to brookslr@ssc.vax.mcmaster.ca.

culty resulting from a complicated rule, and marginal difficulty that can be avoided by making special procedures for repetitious particular circumstances.

We focus on a different type of influence that could lead to specialized procedures: episodic retrieval. The cues presented by an item to be processed provide information necessary for applying a general rule, but they also provide retrieval cues for prior episodes involving either that item or similar items. Under the right retrieval conditions, these prior processing episodes, together with associated outcomes and special considerations, can intrude on and modify current processing. Of course, remembering a prior answer or tricky aspect of a particular item can be expected to reduce the difficulty of processing that and similar items. But the process of specialization can profitably be thought of as being strongly influenced by the whole range of factors affecting episodic retrieval. Further, specific retrieval may be strong enough and common enough to change the usual picture in which procedures become more general and abstract with practice.

### Experimental Materials

The imaginary animals shown in Figure 1 divide into two categories: those that live in shelters they build from materials available in their environment (Builders), and those that live in holes they dig (Diggers). The animals can be categorized by a simple (simplest possible) additive rule, involving three of the five binary dimensions used to construct them. We gave subjects a categorization rule that was sufficient to correctly categorize all of the animals in the experiment. The subjects were then trained in categorizing eight sets of these items (e.g., the "known builders" and "known diggers" in Figure 1). Our question is whether this training resulted in a faster application of the rule or (also) established special cases that would influence the categorization of subsequent new items. To determine this, we introduced positive- and negative-match items into the test sequence. According to the rule, a positive-match item is in the same category as the most similar training item. A negative-match item is also similar to one of the training items, but according to the rule, it is in the opposite category to that item. For example, spots and long legs make the negative match of Figure 1 a builder, despite its resemblance to the known digger. If the subjects produced more errors and longer reaction times in the categorization of the negative-match items, we concluded that, for these conditions, the similar past processing episodes had had an effect on the processing of new similar items. The general procedure of categorization would have partly specialized around the particular cases experienced in training, such that new items similar to known items would not be categorized solely on the basis of the original general rule.

### Ecological Considerations

The major part of our argument is not that such effects are likely to occur in general, but only under particular conditions that turn out to be of some ecological interest. We tried to

**RULE:** AT LEAST TWO OF (LONG LEGS, ANGULAR BODY, SPOTS) → BUILDER

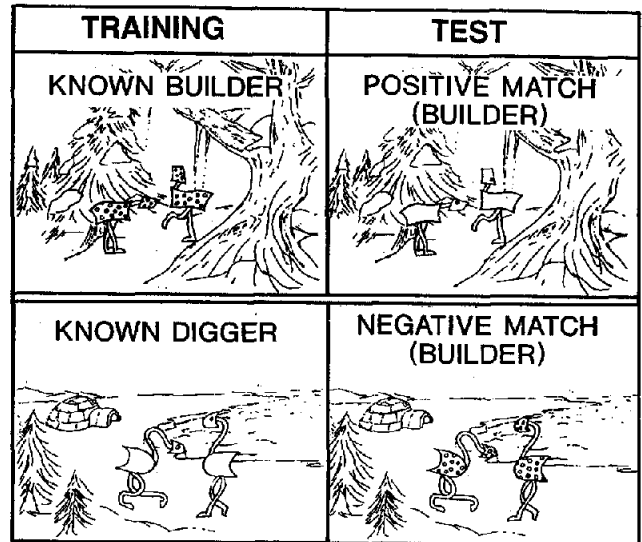


Figure 1. An example of the experimental materials. (According to the rule given, the training items shown are a builder [top left; long legs, angular body, spots] and a digger [bottom left; only long legs]. The positive match appears only in test and, according to the rule, is in the same category as the most similar training item. The negative match is similar to one of the training items but, according to the rule, is in the opposite category to that item [spots in addition to long legs makes it a builder].)

mimic some of these conditions in the experimental conditions that we chose.

1. The stimuli were chosen to have some of the properties of natural objects. Although one can analyze any of the animals, breaking it into its component parts to determine their values, the parts do form an integrated, recognizable whole. This is in contrast to the stimuli often used in concept learning research, which are made up of discrete, nonintegrated dimensions.

2. When we experience an exemplar of a category in "real life" we generally interact with that exemplar on an individual basis. We usually have knowledge specific to the individual item we are dealing with, and we generally encounter the item repeatedly in the same surroundings, interacting with those surroundings. This is accomplished in some of the current experiments by having the subjects learn idiosyncratic information about how each of the training stimuli interact with their environment. Each of the animals builds or digs in a unique way, using materials or techniques appropriate to the environment in which it occurs. The animals occur in different background scenes, and each background is home to one builder and one digger. Although the background scenes do not predict the categories, they do predict the individual animals. Thus, the subjects have information about the individual items and their relative likelihood of occurrence in different surroundings.

3. The form of the rule (see Figure 1) is at least similar to the type of rule often used in verbal communications to

beginners regarding a new visual discrimination. For example, in distinguishing butterflies from moths, one could use a rule of thumb such as "butterflies fly in the daytime, do not emerge from cocoons, have knobs on the ends of their antennae, and have a large wing-to-body-size ratio; moths are generally opposite on each of these dimensions." No one of these features is invariable, but collectively they have a high probability of resulting in correct categorization (even in Africa, where there are some beautiful and rather svelte moths). Comparable rules, which require learning visual categorizations, are given in areas of medicine such as dermatology, histological pathology, radiology, endoscopy, and ophthalmology. In addition, this "additive rule of thumb" is a verbal form that could generate the family resemblance structures that have been held useful for verbally characterizing ill-defined, natural-kinds categories. There is no reason to believe that such rules of thumb are sufficient for actual classification, but they exist, and apparently they are useful instructions for beginners (which is the situation modeled in this article).

4. Finally, and perhaps most important, the test items are embedded in a generally familiar sequence. With the exception of excursions to the zoo or the museum, few areas of life have the kind of concentrated, direct confrontation with strange exemplars that occur in many concept-learning experiments. Previous knowledge resources may be used differently if the subjects' problem-solving abilities are being challenged than if things seem relatively normal.

The results show that not all the conditions just noted are essential to obtain effects of prior episodes, although they all do have at least an influence. Collectively, they do define an interesting set of circumstances that are sufficiently common and sufficiently likely to interact with presumptions about the effects of practice that they deserve investigation.

We found that subjects took longer and made more errors categorizing the negative matches than the other items in the experiments. This we interpret as indicating that similarity to old items influenced categorization despite the subjects having a sufficient and easy categorization rule. A comparable effect was not found when the material was perceptually analytic, such as the material shown in the bottom of Figure 3; that is, the items apparently had to be perceptually and mnemonically distinctive to produce the specialization. The influence of similar prior items was much stronger when testing occurred in a context of "friends"—items that were familiar from the training set—than when testing occurred among items that were novel recombinations of old elements. This suggests that the retrieval context is important in producing intrusions from past episodes. The insistence of these intrusions was suggested by the finding that a specific injunction to watch out for negative-match items was not sufficient to prevent the increase in errors and reaction times on negative-match items. In addition, people generally did not have insight into the source of their errors on negative-match items. These effects of similar past episodes are produced by conditions that are ecologically common and that are of special interest when the categorization rule becomes uncertain, as for example when the rule was presented as being merely a heuristic used for instructional purposes.

## Experiment 1: Initial Demonstration

This experiment was designed to demonstrate the effect of prior episodes on classification performance in the presence of a simple sufficient rule. The rule group was given a completely adequate rule at the beginning of the session; the no-rule group was not. Both groups received the same initial training session, in which they were given feedback about the correctness of their categorizations. If the rule group is influenced by similarity to specific training items, then in the subsequent test phase, they should make more errors on negative-match items than they make on positive-match items, despite the fact that the rule applies equally well to both. If the no-rule subjects respond on the basis of similarity to old items rather than to an induced rule, we would expect a high error rate on the negative-match items, because they would have no other basis of response. This result would provide internal validation of our suppositions about the effect of similarity-based responding. To the extent that these no-rule subjects induced a rule during training, they would behave like the rule group (and consequently would be less useful for our purposes).

Our interest was in discovering conditions that make specific memories highly available. The stimuli were drawings presented on distinctive backgrounds, the test items were presented in a list consisting mainly of old items, and the subjects were required to learn some idiosyncratic information about each of the training items. Subsequent experiments will evaluate the importance of these conditions.

## Method

*Subjects.* Eighty students from the introductory cognition course at McMaster University participated in the study, for course credit.

*Materials.* The stimuli were line drawings of imaginary animals (also used by Brooks, 1978) made up from 5 two-valued dimensions: body type (angular or curved), spots (present or absent), leg length (long or short), neck length (long or short), and leg number (two or six). The animals were categorized as builders or diggers by a three-feature additive rule that used body type, leg length, and spots as relevant features. An animal was defined as a builder if it had builder values on at least two of the three relevant features; all other animals were deemed as diggers. A logical description of the stimuli is given in the Appendix. As can be seen, the value of each relevant feature was consistent with the classification of the animal in 75% of the cases. The values of the irrelevant features appeared equally often in animals of each category, and thus were in themselves nondiagnostic. Each animal differs from each of the others on at least two attributes, with the exception of its matching item, which differs from the closest item only on the dimension of spots. It is important that the feature that is varied to generate the corresponding items is one of the features relevant to the rule. This makes it possible for items to change categories, yet remain similar to an item in the opposite category. If this were not the case, any transfer to new items might be regarded as simply a failure of discrimination between the old and new items. Making the discriminating feature one of the features relevant to the rule makes this possibility much less likely.

Four different rules were used to counterbalance the item across the conditions. The values associated with builders for each of the four rules were (a) long legs, angular body, and spots; (b) short legs, angular body, and no spots; (c) short legs, curved body, and spots;

and (d) long legs, curved body, and no spots. This manipulation of rules, along with variation in which items appeared in the training set and which, in the test set, ensured that each of the items in the experiment appeared equally often in the following four critical conditions.

1. *Positive match*: An item seen for the first time in the test phase, and which is in the same category as its similar old item.
2. *Negative match*: An item seen for the first time in the test phase, and which is in the opposite category to its similar old item.
3. *Positive olds*: Items seen in the training phase, and for which the similar new item will be a positive match (e.g., upper left panel in Figure 1).
4. *Negative olds*: Items seen in the training phase, and for which the similar new item will be a negative match (e.g., lower left panel in Figure 1).

The items appeared on one of four colored background scenes. Each background in the study set was home to one type of digger and one type of builder so that the backgrounds themselves were not differentially associated with builders or diggers.

*Procedure.* Subjects were tested individually. The materials were all presented by means of a slide projector connected to a microcomputer. The projector and computer were connected through a light-sensitive switch, so that the computer recorded the time elapsed from the projection of a slide to the subject's response (pressing a telegraph key).

There were two conditions with 40 subjects each: the rule condition and the no-rule condition. For both groups there was a training phase and a test phase. The subjects in the rule condition were informed of the rule at the start of the training phase. Those in the no-rule condition were never informed of a rule. They were told that the first time they saw an animal they would have to guess whether it was a builder or a digger, but on subsequent trials they would be able to remember what it was.

In the training phase the subjects were given 40 trials made up of the eight old items seen on five trials each, in random order, with the restriction that the same item would not appear twice in a row. Each training trial consisted of three slides. The first slide showed a pair of animals simply standing in a given background (as in Figure 1). Subjects were instructed to classify the animals as quickly as possible without sacrificing accuracy. After their response they were given feedback, and if it was the first presentation of a particular animal, they were shown the next two slides in the set. These second and third slides showed the particular way each animal built or dug its home. On all subsequent presentations of each animal in the training phase, the subjects were again required to make a classification decision based upon seeing the first slide. Before seeing the two "story" slides again, subjects were asked to remember how the animal built or dug. Presenting the second and third slides and requiring recall of information from the subjects was intended to individuate the animals and to prevent them from being processed solely as instances of the rules.

The test set was identical for both groups. There were 40 items, split into a phase in which the only new items were positive matches and a subsequent phase in which the only new items were negative matches. This separation into a positive phase and a negative phase was intended to allow us to evaluate a possible generalized caution effect when the subject discovered the presence of negative matches. The positive phase contained the four positive old items, the four positive-match items, and four repetitions of each of the negative old items. The four repetitions of the negative old items served as fillers, holding the positive-old and positive-match items apart in the series, as well as increasing the ratio of old to new items. The negative phase contained three further repetitions of each of the positive-old items

shown in the positive phase, as fillers, and the four negative-match items. There were a total of 32 old and 8 new items in the test set as a whole. The negative phase immediately followed the positive phase, with no break between the two. In both phases, the filler items were those old items that were least similar to the match items, so that there would be less chance for immediate interference or enhancement from the corresponding old item.

The items in the test phase were presented as pairs of slides. The first slide showed only the background on which the upcoming test item would be displayed. The subjects were simply to look at this background and indicate when they were ready for the second slide. The second slide showed the same background with a pair of animals on it. As in training, subjects were instructed to classify the animals as quickly as possible without sacrificing accuracy. They were also told that they might be able to use the first slide to anticipate which items were most likely to appear on the background and that there would be some new items in this phase of the experiment. No feedback was given in the test phase.

## Results

*Analysis.* We predicted elevated response times and error rates for the negative-match items compared with the positive matches. However, a simple comparison between the positive and negative matches is not appropriate because of practice effects that occurred during the test set. Therefore, the analyses were set up as two  $2 \times 2$  analyses of variance, one for the positive phase and one for the negative phase. For each of these analyses, knowledge of the rule was a between-subjects factor (rule vs. no rule), and previous experience with the item (old vs. positive match or old vs. negative match) was a within-subject factor. In the analysis of the positive-phase, the comparison was between the positive-old items and the positive-matches; in analysis of the negative phase, the comparison was between the first negative-phase presentation of the positive-old items and the negative matches.<sup>1</sup> Only those positive items (positive olds and positive matches) that had two consistent features (either one or two builder features) were included in the analyses. This was done to avoid potential confounding due to half the positive items having all three features that are consistently builder or digger features, whereas this is true for none of the negative items (negative olds and negative matches). This more conservative test was used, although the findings were the same whether only the items with two consistent features or all the items were included in the analysis.

The mean correct response time for each item type was calculated for each subject. Response time analyses were performed on the logarithms of these times, to better comply with the assumption of normal distributions required for the analysis of variance. For the negative-match items for the no-

<sup>1</sup> This negative-phase comparison was used to generate the most conservative test. Comparison of the negative matches with an additional presentation of the negative-old items in the negative phase would have been inappropriate due to use of the negative-old items as fillers in the positive phase, and the consequent practice effects on those items. Similarly, comparison of the negative matches with the first presentation of the negative-old items in the positive phase was inappropriate because of generalized practice effects.

rule group, the response times were calculated for incorrect rather than correct responses. For these subjects, who were not aware of the rule, wrong answers on the negative-match items are actually the most appropriate responses. Whether correct or incorrect answers are used for this cell turns out not to be critical to the results, because whether the mean correct (1,149 ms) or incorrect (1,120 ms) response time is used, the difference between the old and match items for the no-rule condition is, as expected, in the opposite direction than that for the rule condition. The criterion alpha level was set to .05 for all analyses.

**Errors.** The error data are shown in Figure 2. There were no significant differences between any of the items in the positive phase. In the negative phase there were more errors for the no-rule group than for the rule group,  $F(1, 78) = 27.8$ ,  $MS_e = .05$ ; more errors for the matching items than for the old items,  $F(1, 78) = 192.3$ ,  $MS_e = .06$ ; and an interaction between rule presence and previous experience such that there was a larger difference between the old and matching items

in the no-rule condition than in the rule condition,  $F(1, 78) = 34.8$ ,  $MS_e = .06$ .

**Response times.** Seven subjects, 6 from the rule group and 1 from the no-rule group, were dropped from the response-time analysis because of empty cells.<sup>2</sup> The response time data for correct responses are shown in Figure 2. As with the error data there were no differences in the positive phase (all  $F$ s < 1). In the negative phase, times were longer for the rule group,  $F(1, 71) = 4.6$ ,  $MS_e = .319$ ; longer for the matching items,  $F(1, 71) = 5.5$ ,  $MS_e = .08$ ; and there was an interaction between rule presence and previous experience such that the match items took longer to classify than the old items in the rule condition, whereas the opposite was true in the no-rule condition  $F(1, 71) = 8.0$ ,  $MS_e = .08$ . The mean response times for the incorrect negative-match items, the only items for which enough errors were made to obtain a reasonable estimate of response times, were 1,172 ms for the rule group and 1,120 ms for the no-rule group (see previous discussion of response times for the means of the no-rule group negative matches).

## EXPERIMENT 1 ERROR RATES

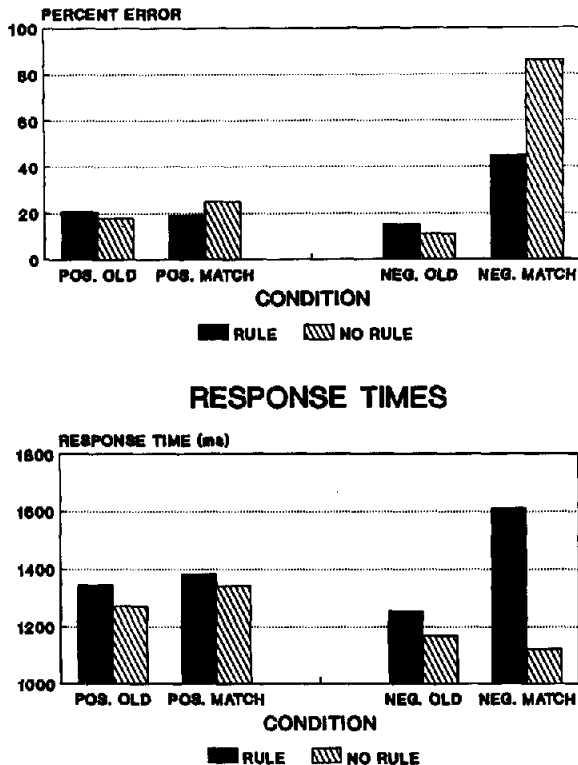


Figure 2. The results for Experiment 1. (For the group that knew the rule, both errors and response times were high for the negative-match items but were unaffected for the positive-match items relative to the old [training] items. The rule applied equally well to both the positive- and negative-match items. Although the no-rule group did not know the rule, and therefore were not technically making errors, the results are plotted as percentage of responses inconsistent with the rule [for easier comparison with the rule group].)

## Discussion

The main result of this experiment for both the rule and the no-rule groups is the large increase in errors on the negative-match items but not on the positive-match items. This difference is consistent with the uses of similarity to previously seen items as a classification strategy. This difference would not be expected if subjects were simply using the rule they were given or, in the case of the no-rule subjects, if they were relying on a rule that they had induced during training. On the other hand, both the accuracy and the speed results suggest that the subjects were not just ignoring the rules and simply using a similarity strategy. The no-rule group made 86% errors on the negative-match items (or 14% responses consistent with the rule) as opposed to only 45% errors made by the rule group. For the response times, there was a substantial slowing of the responses for the negative matches only for the rule group, suggesting that for these subjects the rule-based response and the similarity-based response were in conflict. This increase in response times would not be expected in the no-rule condition, assuming that a rule was not induced to conflict with the similarity-based response.

This reliance on similarity to specific prior episodes occurs in a situation that does not bias against the use of the rule by withholding the rule from the subjects or by supplying them with an extremely complex rule. There was no necessity in this experiment for the subjects to use anything but the rules they were given. The rules were perfectly predictive, yet the subjects made significant errors when new items were similar to previously seen items from the opposite category. However,

<sup>2</sup> These were subjects who failed to correctly classify any of the items in one of the cells. In the rule group 3 subjects failed to correctly classify a negative match and 1 subject each failed to classify a positive-old item, a negative-old item, and a positive match. One subject from the no-rule group failed to correctly classify a positive-old item.

as noted in the introduction, this experiment was designed to facilitate the availability of specific instances. Subjects had idiosyncratic information in addition to the classification of an item (the story slides). The test was conducted in a situation where the subjects were familiar with most of the items they were seeing (80% old items), and the context was established before the item was shown (the background slides). We believe that, in a number of ways, such conditions mimic one kind of real-world situation. Nonetheless, we should explore the importance of these variables. The following experiments will explore some of the limiting conditions of the similarity-based specialization of the application of explicit rules demonstrated in this experiment.

### Experiment 2: The Effect of Perceptual Integration

Experiment 1 demonstrated that applying a rule to a restricted set of instances can have the consequence of establishing item-specific bases of generalization. However, there is no reason to presume that practice with a small number of exemplars always has this effect. Intuitively, when an explicit rule is practiced, the learner often simply becomes faster at applying the rule. One set of variables that could be expected to control these different outcomes of practice is the perceptual and mnemonic characteristics of the exemplars. For example, the list of features shown in Figure 3 is logically equivalent to the explicit dimensional structure of the animals used in Experiment 1 and would allow the application of identical classification rules. Presenting the instances in this form might not lead to integration of the information into mnemonically separate instances. Rather, subjects might have a greater tendency to treat the display as separable pieces of information and not learn them incidentally as special cases. If this happened, then we would not expect practice with the feature lists to result in differences in performance between positive and negative matches, nor, for that matter, between new and old items.

The second contribution of this experiment is to show the robustness of the basic results of Experiment 1. In Experiment 2, the story slides were not appropriate to the feature lists; thus, they were not included for the drawn animals either. Further, Experiment 2 includes fewer old items in the test phases. Finally, the context for individual items in the transfer set was not established by initial exposure to background-only slides. Incidentally then, this experiment will demonstrate that the story slides of Experiment 1, the large proportion of old items during the test phase and the background-only slides used during the test phase are not essential to producing the similarity effects with the animal stimuli.

### Method

**Subjects.** The subjects were 72 McMaster University undergraduate summer students, who were paid for their participation.

**Materials.** For the drawing group, the stimuli were the same stimuli used in Experiment 1. For the feature list group, the stimuli were logically equivalent lists of features rather than drawings. The features were always listed in the same order. The training set con-

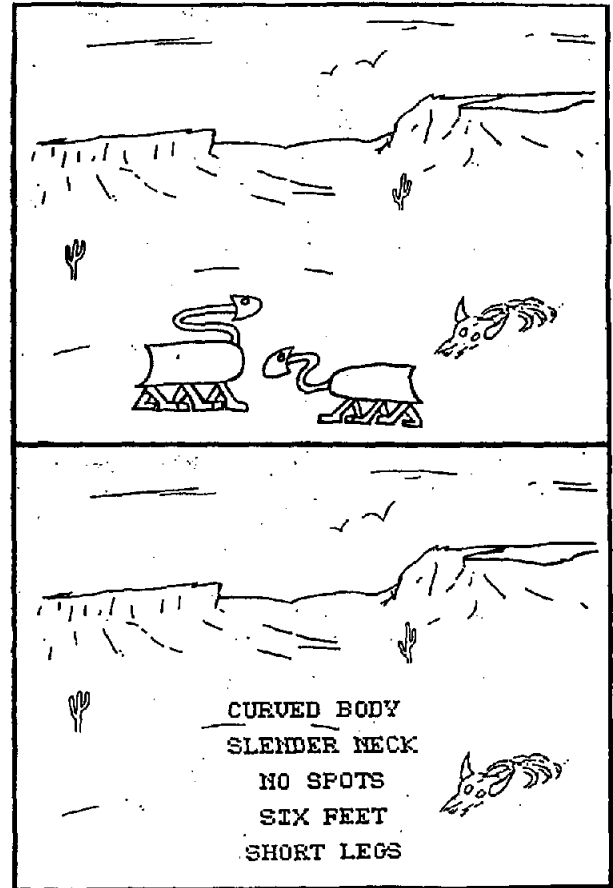


Figure 3. An example of the materials used in Experiment 2. (The drawing group saw stimulus items similar to those used in Experiment 1. The feature-list group saw items in which the same sets of features were presented as verbal descriptions rather than as pictures.)

sisted of eight items, as in the first experiment, however in Experiment 2 only one rule was used to define builders and diggers.

The test set contained more types of items in this experiment than in Experiment 1. These differences will be listed but not expanded on here, because they are not important to the present discussion. In addition to varying as old items and positive and negative matches, the animals in the test phases could be presented on appropriate or inappropriate backgrounds. An appropriate background was the background on which an animal, or its matching old item, had been seen in the training set. An inappropriate background would be one of the other three backgrounds. The test set was broken into two parts. The first part of the test set contained the following items, all presented in random order: the eight old items on the appropriate backgrounds, the four positive matches on appropriate backgrounds, the four positive-old items on inappropriate backgrounds, and the four positive matches on inappropriate backgrounds. The second part of the test set consisted of a second presentation of the eight old items on the appropriate backgrounds, a second presentation of the four positive matches on appropriate backgrounds, a second presentation of the four positive matches on inappropriate backgrounds, the first presentation of the four negative matches on appropriate backgrounds, and the first presentation of the four negative matches on inappropriate backgrounds. Thus, across the two test sets there were

16 old (a previously encountered animal on the previously encountered background) and 28 new items. The items used in the present analyses are the old items, the positive-match items, and the negative-match items, all on the appropriate backgrounds.

*Procedure.* The subject was first presented with a set of study items, which consisted of the eight old items. There were five repetitions of the eight items, presented in five randomly ordered blocks, for a total of 40 trials. The subject was informed of the rule and asked to classify each item by pressing one of the two response keys as quickly as possible without sacrificing accuracy. Feedback was given after each item. Then a new slide tray, containing the first test set, was loaded, and the subject was instructed to continue doing the same task, except that there would be no more feedback. When the first test set was completed, another slide tray, containing the second test set, was loaded, and again the subject was instructed to continue with the same task.

## Results

*Analysis.* The analyses for this experiment are similar to those for Experiment 1. There were two  $2 \times 2$  mixed ANOVAs, one for the positive items and one for the negative items. The between-subjects factor was the form of the stimulus representation (animals vs. feature lists), and the within-subject factor was experience with the item (old vs. positive or negative match). In all other respects the analysis was identical to that for Experiment 1. Two subjects, one each from the drawing and feature-list conditions, were dropped from the analysis of the response times because of missing data.

*Errors.* The error data are shown in Table 1. For the positive phase, there were no significant effects due to the form of the stimuli or due to previous experience or their interaction (all  $F_s < 1$ ). For the negative phase, the drawing group made more errors than the feature-list group,  $F(1, 70) = 8.42$ ,  $MS_e = 249.5$ ; the old items showed fewer errors than the negative-match items,  $F(1, 70) = 17.46$ ,  $MS_e = 223.7$ ; and there was an interaction between stimulus representation and previous experience,  $F(1, 70) = 13.12$ ,  $MS_e = 223.71$ . Thus, there was an increase in error rates for the negative matches only for the drawing group.

*Response times.* The response time data are shown in Table 1.<sup>3</sup> For the positive phase the drawing group showed faster responses than the feature-list group,  $F(1, 68) = 27.75$ ,  $MS_e = 1.61$ ; old items showed faster responses than positive-match items,  $F(1, 68) = 13.15$ ,  $MS_e = .033$ ; and there was no representation-experience interaction,  $F(1, 68) = 1.86$ ,  $MS_e =$

0.33, not significant. For the negative phase the drawing group was again faster than the feature-list group,  $F(1, 68) = 5.25$ ,  $MS_e = .164$ ; old items were classified faster than negative matches,  $F(1, 68) = 16.33$ ,  $MS_e = .029$ ; and there was an interaction between stimulus representation and previous experience,  $F(1, 68) = 16.07$ ,  $MS_e = .029$ . Thus, there was an increase in response times for the negative matches only for the drawing group. The incorrect response times for negative matches were 1,224 ms for the drawing group and 2,175 ms for the feature-list group (although the mean for the feature-list group should be viewed with some suspicion because it is based on so few observations).

An additional  $2 \times 2$  within-subjects ANOVA on response times was performed on the old items from the feature-list group only where the factors were test (first or second) and status of matching item (positive or negative). This analysis compared the response times for the first presentation of the old items during test (1,887 ms for the positive olds and 1,971 ms for the negative olds) with the times for their second presentation (1,751 ms for the positive olds and 1,753 ms for the negative olds). The only significant effect was that of test, where the first test showed slower response times than the second,  $F(1, 34) = 12.33$ ,  $MS_e = .02$ , suggesting that the subjects were getting faster at answering in the feature-list condition even though they showed no signs of nonanalytic specialization (no difference in performance on positive and negative matches).

## Discussion

Subjects who saw drawings of the animals in both training and test showed evidence of instance-based specialization of the rule, namely, increased errors and response times for the negative-match items. Subjects who had an equivalent number of exposures to feature lists in both training and test did not show this evidence of instance-based specialization. For these subjects the effect of practice was simply to speed up the rule application. These results suggest that evidence of instance-based classification is more likely with stimuli that can easily be integrated into mnemonically distinct instances. The drawings subjects were able to respond much more rapidly with an equivalent number of training trials, but apparently at the cost of greater vulnerability to negative matches when they occurred. We suspect that specialization effects can be shown with verbal stimuli; the work of Jacoby (e.g., 1983b) and of Whittlesea (1987; Whittlesea & Cantwell,

Table 1  
Mean Responses in Experiment 2

Group	Positive phase		Negative phase	
	Old	Match	Old	Match
	% errors			
Drawings	4.2	2.8	4.9	24.3
Feature lists	4.2	4.2	6.3	7.6
	Response times (msec)			
Drawings	1,192	1,413	1,368	1,691
Feature lists	1,750	1,881	1,752	1,751

<sup>3</sup> Neither the Test 1 items nor the inappropriate items from Test 2 were analyzed. For the drawings group, the mean percentage errors for these items were 6.9, 5.6, 6.9, 5.6, 1.4, 5.6, and 25 for the negative-old appropriate, positive-old appropriate, positive match appropriate, positive-old inappropriate, positive match inappropriate, second-presentation positive match inappropriate, and negative match inappropriate items, respectively. The corresponding mean correct response times (in milliseconds) were 1,303, 1,355, 1,382, 1,599, 1,545, 1,422, and 1,787. The corresponding means for the feature list group were 6.3, 8.3, 6.9, 8.3, 5.6, 4.2, and 4.2 for errors and 1,971, 1,887, 1,977, 1,862, 1,905, 1,742, and 1,921 (milliseconds) for response times.

1987; Whittlesea & Brooks, 1988) certainly suggest that they can. But at the least, there would have to be categorically irrelevant variation in the typefaces, order, type of prose, or processing task to provide sufficient differentiation.

In addition, the similarity between the results of the drawing condition in this experiment and of the rule condition in Experiment 1 suggests a few things that are not in themselves critical to the occurrence of instance-based specialization of the rule. These are information in training other than category membership (the story slides in the training set), establishment of the context before the appearance of the animal (the background-only slides in the test set), and the very high proportion of old items in Experiment 1 (32 of 40 in Experiment 1 as compared with 16 of 44 in Experiment 2). The next experiment and the subsequent reanalysis will demonstrate that the similarity effects are robust across other potentially limiting features of the initial demonstration.

### Experiment 3: Caution and Awareness of Error Source

In both previous experiments, subjects were encouraged to respond as rapidly as was consistent with accuracy. It is possible that the intrusion of specific familiarity on rule-governed classification is limited to conditions with this emphasis on speed. The purpose of the third experiment was to change the emphasis of the subjects increasingly toward accuracy, to see where the effects of negative matches disappear. The first of three instructional groups, the speed group, had identical instructions to those used in Experiments 1 and 2. With the exception of a list context manipulation that will be described in the method section, this group was a replication of the rule condition of Experiment 1. The accuracy group was told that accurate classification was to be their main concern. They were given feedback on any errors in the test phase, with a reemphasis of the need for accuracy. Finally, the alert (to negative matches) group, in addition to having accuracy instruction and feedback, was told of the presence of negative matches and had the first negative match pointed out to them immediately after they classified it. In the previous two experiments, subjects frequently gave immediate comments on any errors that they made, even though they were not given specific feedback during the test phase. However, without the manipulations of Experiment 3, it cannot be clear whether this realization that a mistake had been made would be translated by the subjects into an emphasis on accuracy, or that they realized the source of their mistakes on negative matches.

### Method

*Subjects.* The subjects were 120 students from the introductory cognition course at McMaster University. They participated for course credit.

*Materials.* The materials were exactly those used in Experiment 1 except for the items used as fillers in the test set. The makeup of the training set was the same for all three conditions and was identical to the training set for the rule group in Experiment 1. The test set

was also identical to the test set in Experiment 1, with the exception that the filler items were all seen on inappropriate backgrounds. In Experiment 1, all items were always seen on the appropriate background, that is, the background on which that item or its matching old item was seen in training. For reasons that will be discussed in the subsequent section, we wanted to decrease the proportion of precisely old items in the test. To accomplish this, all the filler items in the present experiment were seen on inappropriate backgrounds, so that the only items seen on the appropriate backgrounds were the critical items. Thus, the positive phase of the test set was made up of the four positive-old items, the four positive-match items, and four repetitions of each of the negative-old items seen on a different inappropriate background for each repetition (except for the fourth presentation on which the first inappropriate background was repeated). The negative phase of the test set was made up of the four negative-match items, along with three repetitions of each of the positive-old items, seen on a different inappropriate background each time. It is worth repeating that the old items are not themselves positive or negative with respect to the rule, but only that their matching new item would, according to the rule, be in the same or opposite category.

*Procedure.* There were three conditions in the experiment, with 40 subjects participating in each group. The groups differed in the instructions they received before being given the training set and again before the test set. The speed group was given the same instructions as the rule group from Experiment 1, that is, to classify the items as quickly as possible without sacrificing accuracy. The accuracy group was given instructions that emphasized accuracy. They were told that response times were also being measured, but that this was of secondary importance and that their main concern should be with accuracy. These subjects also received feedback regarding the accuracy of their responses in the test phase, again emphasizing the importance of accuracy. The instructions to the subjects in the alert condition were identical to those in the accuracy condition, except that they were alerted to the presence of negative-match items. They were told at the beginning of the test sequence that there would be some new items in the test set that closely resembled items they had seen before, but that would in fact be in the opposite category, and that they should watch out for these items. In addition, when the first negative match was presented, the experimenter pointed out to the subject that it was one of the "tricky" items. The experimenter told the subject which old item it was similar to, pointed out that it was in the opposite category, and again warned the subject to watch out for this type of item.

### Results

The analysis of the results of this experiment, like the materials and design, was very similar to that of Experiment 1. The data were subjected to two  $3 \times 2$  mixed ANOVAs, one for the positive phase and one for the negative phase, in which the between-subjects factor was the instructional condition (speed vs. accuracy vs. alert) and the within-subject factor was item type (old vs. match). Two subjects from each of the speed and accuracy groups and 1 from the alert group had to be dropped from the response time analyses because of empty cells. All other aspects of data analysis were as in Experiment 1.

*Errors.* The error data are shown in Table 2. In the positive phase, there was a marginal interaction between the instructions and item type,  $F(1, 117) = 3.02$ ,  $MS_e = .05$ ,  $.05 < p < .10$ . Neither main effect approached significance. In



Table 2  
Mean Responses in Experiment 3

Group	Positive phase		Negative phase	
	Old	Match	Old	Match
	% errors			
Speed	13	13	6	28
Accuracy	5	19	8	19
Alert	9	8	4	18
	Response times (msec)			
Speed	1,404	1,609	1,345	1,348
Accuracy	1,469	1,478	1,352	1,519
Alert	2,012	2,157	2,003	1,936

the negative phase, there was no main effect of the instructions, nor was there an Instruction  $\times$  Item-type interaction. There was, however, a significant effect of item type, with the subjects showing more errors on the negative-match items than on the negative-old items,  $F(1, 117) = 35.14$ ,  $MS_e = .04$ .

*Response times.* The response time data are displayed in Table 2. In the positive phase there was a significant main effect of instructions,  $F(2, 112) = 11.88$ ,  $MS_e = .25$ , as well as a main effect of item type,  $F(1, 112) = 5.11$ ,  $MS_e = .07$ . There was no interaction between the two. In the negative phase there was a main effect of instructions,  $F(2, 112) = 17.49$ ,  $MS_e = .19$ , but the item-type and interaction effects were not significant. The incorrect response times for the negative-match items were 1,254, 1,146, and 1,535 ms for the speed, accuracy, and alert conditions, respectively.

### Discussion

The increase in errors on the negative-match items in both the accuracy and alert groups suggests that the instance-based effects in the speed group (and the comparable groups in Experiments 1 and 2) were not entirely due to subjects' emphasis on speed. When subjects were asked to concentrate on accuracy and were given feedback to point out their errors, there was little change from the results under the "speed without sacrificing accuracy" instructions given to the speed group and to subjects in the previous studies. When subjects were alerted to the nature of the difficult items they would encounter, they decreased their speed considerably but gained little in their accuracy, and they continued to show increased errors on the negative-match items. This pattern of results suggests that an effect of specific training items intrudes, despite at least some effort on the part of the subject to prevent it. This is especially true in the alert condition in which the subjects should have been alerted to the problem, but still seemed unable to prevent errors.

The high error rate for positive matches in the accuracy condition is problematic. This was unexpected and does not fit with either a rule-based or a similarity-based strategy. This is the only time we have seen such a difference between old and positive-match items. The increase in error rates for these items is accompanied by a lesser increase in response times than in the speed or alert conditions and may represent a

speed-accuracy tradeoff, favoring speed for these items. This is not a particularly satisfying explanation, however, because the tradeoff is in the opposite direction to what the instructions required. However, although there is an approximately equal increase in errors on both positive and negative matches in the accuracy condition, it seems clear that, overall, there is a difference in the two types of items; the increase in errors on positive matches disappears in the alert condition, but the increase on the negative matches does not.

Additional evidence from the accuracy group suggests that the similarity information intrudes without the subject even being particularly aware of it. In a postexperimental interview, 24 subjects from the accuracy condition were questioned regarding which items gave them problems and why they thought they might be faster on some items than on others. They were shown pairs of items consisting of one positive match and one negative match and told they answered faster on the positive match. They were asked why they thought this might be the case. Only 10 of the 24 subjects ever mentioned similarity to an old item in the opposite category as a possible reason for the relatively poor performance on the negative matches. Furthermore, those subjects who were counted as mentioning similarity were those who made any mention of similarity at any point in the discussion, often after considerable prompting from the experimenter. Thus, even using the most generous of criteria, over half of the subjects failed to attribute their poor performance on those items to the fact that they were negative matches. They generally picked features of the animal that seemed to be plausible causes. For example, "this one had a long neck, so I must have thought it had long legs and called it a builder" and "this one had long curvy legs, and so I guess I got them mixed up with its body, which is angular, and called it a builder by mistake". Occasionally, the subject could not find a plausible reason in the slide and said, "I guess I was trying to answer too fast," which of course is not a reason for relative decrements in performance. The subjects' ignorance of the source of their errors, as well as their inability to reduce their errors on negative matches when alerted, suggests that their use of similarity to old items is not necessarily under deliberate control.

### Reanalysis: Familiarity of List Context

One of the ecological constraints described in the introduction was a test context that contained a predominance of familiar items. In this section, two groups from previous experiments will be compared to evaluate the effect of this variable. The rule group from Experiment 1 and the speed group from Experiment 3 differed only in the relative familiarity of the filler items. The rule group from Experiment 1, referred to below as "friends list context," used old items as fillers. This resulted in 32 of the 40 items in the test list being items seen in exactly the same form as in training. The filler test items for the speed group from Experiment 3, referred to below as "strangers list context," consisted of old animals appearing on a variety of different backgrounds (a different background for each trial in which a given filler animal

appeared). This re-pairing of animals and backgrounds was designed to provide a list context that was less familiar (4 old items, 8 new matching items, and 28 re-paired items) than with the friends list context. This changed list context could make the subject rely on the rule more, or possibly even make the training items less available, analogous to the mnemonic deficit resulting from changing list context in memory experiments (a comparable effect of familiar list context was found for the perceptual identification of words by Jacoby, 1983b). Although the list contexts (filler items) differed between the two test lists, the analyses below (and in the previous experiments) are based only on those 12 critical items that were identical in the two lists.

### Method

**Analysis.** The data are presented in Table 3. The analyses consisted of  $2 \times 2$  mixed ANOVAs with list context (friends vs. strangers) as the between-subjects factor, and item type (old vs. match) as the within-subject factor. The criterion alpha level was set to .05 for all analyses.

**Errors.** In the positive phase, there was a marginal effect of the list context, with more errors in the friends condition,  $F(1, 73) = 3.13$ ,  $MS_e = .07$ ,  $.05 < p < .10$ , and no effect of item type or the list context by item-type interaction (both  $F$ s  $< 1$ ). So that, although there was a trend toward an advantage for the strangers group, there was no advantage for old items or interaction between list context and item type. In the negative phase, there was a significant effect of list context,  $F(1, 78) = 11.15$ ,  $MS_e = .06$ , and a significant effect of item type,  $F(1, 78) = 50.59$ ,  $MS_e = .05$ , but no interaction between the two,  $F(1, 78) = 1.24$ ,  $MS_e = .05$ , not significant. Thus, in the negative phase, there was an advantage for the strangers group and an advantage for old items.

**Response times.** In the positive phase, there was no overall effect of either the list context,  $F(1, 70) = 1.56$ ,  $MS_e = .35$ , not significant, or item type,  $F(1, 70) = 1.48$ ,  $MS_e = .09$ , not significant, while their interaction was marginally significant,  $F(1, 70) = 3.23$ ,  $MS_e = .09$ ,  $.05 < p < .10$ . In the negative phase there was again no overall effect of list context ( $F < 1$ ), but there was a main effect of item type,  $F(1, 70) = 6.64$ ,  $MS_e = .08$ , and an interaction,  $F(1, 70) = 6.01$ ,  $MS_e = .08$ . Thus, in the negative phase, subjects responded more slowly on the matching items. This was due to longer response times for the negative matches in the friends condition but not in the strangers condition.

### Discussion

Clearly the list context during test did influence instance-based performance. The negative matches in the friends list

context showed many more errors than in the strangers list context. However, the basis of this effect is not as clear. The simplest result to support the hypothesis that list context affects the *availability* of prior exemplars would be one in which olds and positive matches are recognized faster and more accurately and negative matches are more disrupting in the friends condition. This increased disruption did happen for the negative matches, but the results for the olds and the positive matches showed signs of a speed-accuracy tradeoff. Negative olds, positive olds, and positive matches were all in the direction of more errors and faster times in the friends than in the strangers condition. A post-hoc analysis of only the old items showed a significant effect of list context, with subjects making more errors in the friends list context,  $F(1, 78) = 5.12$ ,  $MS_e = .06$ , and a marginal effect of test phase, with subjects performing better in the negative phase than in the positive phase,  $F(1, 78) = 3.17$ ,  $MS_e = .05$ ,  $.05 < p < .10$ . The interaction did not approach significance, and a similar analysis on the response times showed no significant effects. Even though the planned comparisons did not reach significance, the direction of the differences was consistently in the opposite direction to that predicted by a simple effect of increased exemplar availability, a reversal that should not be ignored. However, even if there were a lower criterion for accuracy, this would not exclude an effect based on instance availability. It would not be unreasonable for a more familiar list context to both result in relaxed criteria (leading to somewhat faster and less accurate responses) and provide more available prior instances. Regardless of the basis, what is clear is that the familiarity of the context is important in determining the relative contributions of rule-based and of instance-based generalization.

### Experiment 4: Dealing with a Suspect Rule

In all the experiments so far the rule has had absolute authority. It has been perfectly predictive and was presented to the subjects as such. But in many natural situations rules are substantially less well-established. A rule may be presented as only a rule of thumb, or may be a rule that a learner is in the process of developing. When such a provisional rule starts producing errors, memory for prior instances takes on a special status. Under most natural circumstances, when errors on new items suggest the need to modify a rule, it is unlikely that the responses to known items should be changed. In addition, if a new item looks sufficiently similar to a known item, classifying the new item on the basis of its similarity to the old item, rather than on the basis of the suspect rule, may be worth considering. That is, if the two items match on many characteristics and there is uncertainty about which characteristics are relevant, then it is reasonable to guess that the few nonmatches are not sufficiently important to change the categorization. With a suspect rule, the use of a literal similarity strategy may be advantageous for both generation of better classification performance during a transition period and use as evidence in attempting to frame a better rule.

In the fourth experiment, the rule was originally presented to the subjects as a rule of thumb. The rule would be correct

Table 3  
Mean Responses in Reanalysis

Group	Positive phase		Negative phase	
	Old	Match	Old	Match
Friends	12	19	15	45
Strangers	13	13	6	28
Response times (msec)				
Friends	1,345	1,385	1,253	1,612
Strangers	1,404	1,609	1,345	1,348

most of the time, but there would be a few exceptions. In no case would the classification of a previously encountered item change. The exception items were produced by designating the correct response for the negative-match items as the response associated with the similar prior item (the negative olds). Thus the exceptions would be classified incorrectly if the rule were used, but correctly if an appeal to the most similar old item were used. The nonexceptions should be classified correctly in either case. This feedback rule essentially reifies the ecological argument, given in the preceding paragraph, regarding the use of similarity strategy.

Classification on the basis of similarity to old items should work better if the old items are more memorable and if overall similarity is more salient. Without mnemonic salience of instances, the learner may be more likely to respond to exceptions by changing the features specified in the rule. As in Experiment 2, the drawings were set against the feature lists, with the prediction that there would be more evidence of similarity-based decisions for the drawings. With the feature lists, subjects would be more likely to alter their classification policy in such a way as to inadvertently reclassify some of the old items.

## Method

*Subjects.* The subjects were 24 students from the introductory cognition course at McMaster University, who participated for course credit. All subjects were tested individually.

*Material.* The stimuli used were the same as in all the previous experiments. Both the drawing and feature lists were used. Once again the set of four different rules was used across subjects to avoid confounding the particular items with the logical type of item they represented.

*Procedure.* The subjects were divided into a feature-list group and a drawing group. Throughout the experiment, the feature-list group saw the feature-list descriptions of the animals described in Experiment 2. The drawing group saw the drawings throughout the experiment. Both groups were instructed to respond as rapidly as was consistent with accuracy.

The first phase of the experiment was designed to familiarize the subjects with the rule, and with some of the animals. Subjects were told the rule, and then given a set of eight animals to classify. For the feature-list group the task was to classify each slide as a builder or digger, with feedback being given after each slide. For the drawing group the procedure was the same as the procedure in the training phase of Experiment 1. At the end of the familiarization phase, all subjects had seen eight items five times each, and in addition, the drawing group had idiosyncratic information about the behavior of each animal.

The second phase was the "drift" training phase (the correct rule "drifted" away from the original rule, rather than dramatically changing). At the start of this phase, the subjects were informed that the rule they had been given was a rule of thumb, rather than a perfectly predictive rule. Some of the items they would see in this part would be exceptions to the rule, and their job was to try to improve on the rule. They were then presented with a set of 16 slides, of the following three types: (a) one presentation of each of the eight familiarization slides (olds), (b) two repetitions of each of two new nonexceptions, and (c) two repetitions of each of two new exceptions. With this distribution of items, as in previous experiments, there was a predominance of old items, and most new items were not exceptions. The

new items were generated, as in previous experiments, by changing one relevant feature (the spots). The exception items were those items whose category membership changed when the relevant feature was changed (the negative-matching items in previous experiments). Thus, the exceptions would be classified incorrectly if the rule were used, but correctly if an appeal to the most similar old item were used. The nonexceptions should be classified correctly in either case. The subjects classified the items and were given feedback after each slide.

The third phase was the test phase. Again there were 16 items: (a) one presentation of each of the eight old items, (b) two presentations of each of two new nonexceptions, and (c) two presentations of each of two new exceptions. The new items were each presented, once on the background on which the corresponding old items were seen in the familiarization set (appropriate background), and once on a background that had been seen with different animals (inappropriate background). The order of presentation for background appropriateness was counterbalanced across items, with the restriction that half the new test items preceded their matching old item (with maximum average lag), and half followed them. No evident difference resulted from this balancing, so it was dropped from subsequent analyses. No feedback was given in this section.

## Results

The criterion alpha level was set to .05 for all analyses. The data are summarized in Table 4. Because of the small number of new items, nonparametric analyses were used throughout. For old items, the drawing group showed more correct (rule-based) answers in both the drift-training and test phases, Mann-Whitney  $U(12) = 28.5$  and  $U(12) = 35.5$ , respectively. In the drift-training phase, both groups showed a higher level of rule-based responding for the first presentation of an exception item than for its second presentation. Using a Wilcoxon matched-pairs signed-ranks test, for the drawing group,  $T(12) = 1.5$ ; for the feature-list group,  $T(12) = 10.5$ . However, the level of rule-based responding on the second presentation of the exception items was higher for the feature-list group than the drawing group,  $U(12) = 32.5$ . In the test phase, the feature-list group showed a higher level of rule-based responding than the drawing group on exception items seen on the appropriate background,  $U(12) = 37.5$ , and there was a marginally higher level of rule-based responding within the drawing group on the inappropriate background items than on the appropriate background items,  $T(12) = 19.5$ ,  $.05 < p < .10$ .

## Discussion

In the drift-training phase, the feature-list group and the drawing group showed more correct responses (less rule-based) on the second presentation of the exception items. The subjects evidently remembered the correct response from the first presentation of the exception items and used this to produce better performance on the second presentation. Apparently the drawings were easier to use for this purpose, since the drawing group performed better than the feature-list group on the second presentation (but not the first).

In the test phase there was better responding (less rule-based, more consistent with the classification of matching old

Table 4  
*Mean Proportion Rule-based Responses in Experiment 4*

Group	Drift-training phase				
	Olds	New nonexceptions		New exceptions	
		First presentation	Second presentation	First presentation	Second presentation
Drawing + story	.99	1.0	.96	.88	.33
Feature list	.89	1.0	1.0	.96	.71

Group	Test phase				
	Olds	New nonexceptions		New exceptions	
		Appropriate background	Inappropriate background	Appropriate background	Inappropriate background
Drawing + story	.97	.92	1.0	.54	.75
Feature list	.87	.88	.96	.88	.83

items) for the drawing group than for the feature-list group for the exception items that appeared on the appropriate background. When the exception items appeared on an inappropriate background, this difference was reduced. This indicates that the drawings group made more use of similarity-based responding, where the subjects would override the rule response if a new item was similar enough to a previously seen item. The marginally better performance (less rule-based responses) for the drawing group on the appropriate-background exceptions suggests, as did the previous friends-strangers analysis, that this similarity strategy is dependent upon nominally irrelevant aspects of context.

The difference in performance on old items is also consistent with this interpretation. In both the drift-training phase and the test phase the drawing group performed better than the feature-list group on the old items (more rule-based responses, correct for these items). This would follow if the drawing group could recognize the old items better and would thus be less likely to change their operative rule for these items. Because they have a way of dealing with exceptions (using similarity to old items) and are more likely to recognize old items, there is less pressure on this group to adjust a rule that is correct most of the time. The feature-list group, on the other hand, faced with errors that cannot easily be corrected, had to change their operative rule, and in doing so they risked errors on previously seen but unrecognized items that were correctly classified in the past.

In summary, we have shown that the drawing stimuli, which earlier resulted in more similarity-based responding, led to superior performance under conditions in which the rule is no longer perfectly predictive. Admittedly, the method used here for classifying the exceptions (all the negative matches were to be categorized according to similarity) was designed to generate better performance from a subject using a similarity strategy. However, if we accept the argument (made for example by Rosch, 1978) that many natural categories have a clustered structure not particularly amenable to simple rules, then the type of feedback rule that we used should not be ecologically uncommon. Thus, Experiment 4

demonstrated that the subjects had no particular trouble with such a category structure provided that the stimuli were perceptually and mnemonically distinct. When the stimuli were feature lists, there is no evidence that similarity was used at all.

### General Discussion

The experiments here show that practice with a sufficient and simple rule can lead to application that is specialized around particular exemplars. New items that were similar to old items in the same category were responded to approximately as rapidly as the old items themselves. New items that were similar to old items in the opposite category produced slower responses and a much higher rate of error. By comparison with a group that did not have a rule, it is clear that knowledge of the rule helped to maintain accuracy, but it did not prevent facilitation or interference from item-specific and context-specific knowledge.

The phenomenon seems robust in the face of manipulations that might have forced subjects away from being influenced by past items. Removing additional individuating information in the form of stories (Experiment 2), changing the instructions to emphasize accuracy (Experiment 3), and decreasing the proportion of familiar items in the test series (Reanalysis), all failed to eliminate similarity-based responding, although emphasizing accuracy and decreasing the proportion of old items in the test both decreased the size of the effects. Subjects' failure to avoid the detrimental effects of similarity on the negative-match items in the alert condition of Experiment 3 occurred despite their paying a large premium in response time.

We would like to stress that these results occurred under conditions designed to be of ecological interest, but not under conditions that were maximally biased to produce effects of prior episodes. To the extent that the conditions of the present experiments do not match those of interesting, real world situations, they tend to be biased in the direction of more rather than less rule use. The rule was the simplest possible

additive rule. A fixed classification rule with any less than three binary features does not allow even an approximation to a graded structure. In all the experiments except Experiment 4, the rule was perfectly predictive, an unusual natural situation for a rule this simple. Furthermore, the dimensions were binary, with clearly discriminable values, characteristics that might have led to easier application of the rule than, for example, in many medical settings. Thus, there was no necessity in these experiments for the subjects to use anything but a straightforward rule strategy to perform flawlessly.

### *Conditions for Producing a Specialization Effect*

Our impression is that the intrusion of negative matches is not a fragile phenomenon as long as familiarity remains predominantly useful within the general domain. The combination of old and positive matching items, items for which a retrieved answer would be useful, were preponderant in all experiments (neglecting the variations in background of Experiment 2). Given that the same answer usually resulted from both specific retrieval and rule use, it apparently was hard not to relax into quickly accepting the answer that came from appearances. The distribution of familiar and unfamiliar items is probably also important when considering the potential effect of more extensive practice. The subjects in our experiments obviously had limited practice with the whole task. To make a convincing argument about automaticity, these experiments would have to have been continued for many more trials than they were. However, if additional practice had continued to favor a few old items, as is true with much of our interaction with natural objects, we would expect that specialization effects would continue to occur. Having said this, it is worth repeating that, with the limited amount of practice available in the current experiments, the rule did continue to be used, as evidenced by spontaneous error correction in virtually all cases.

In addition to predominant familiarity, it appears that having a classification rule at least as complicated as an additive rule is important for producing specialization. A simple attributive rule based on a single, perceptually clear feature probably would not have given a specialization effect. However, simple additive rules of thumb seem to be prevalent in instructing adults in new visual categorizations. Initial instruction in medicine and biology contain a sufficient number of examples of simple additive rules for this to be not too restrictive a condition.

Providing perceptually coherent, memorable visual units also seems important. In Experiments 2 and 4, similarity effects were seen when the items were presented to the subjects as drawings, but the effects were not seen when the same information was presented as written lists of features. This is consistent with our emphasis on specific retrieval, because an effect of similar prior episodes requires that the material be processed as distinctive events in memory.

Nosofsky, Clark, and Shin (1989) reported data germane to this issue. They found that when they instructed subjects to use classification rules, the subjects' data were fit better by a

rule model than by an exemplar model, but when subjects were not given a rule, the exemplar model generated the best fit. This contrasts with our finding of strong episodic effects, despite subjects' knowledge of rules. However, the materials that Nosofsky, Clark, and Shin used were circles containing a radius line, with the size of the circle and the angle of the radius varying. Given the nature of the stimuli (two relatively unintegrated features) we would be inclined to compare their results to the results we obtained in the feature list condition. This inclination is borne out by the recognition memory data collected in their study, which show subjects performing at chance levels. This is consistent with our belief that memorable units are important for the occurrence of episodic effects in the presence of an effective rule.

Under processing circumstances different from those used in our Experiments 2 and 4, we believe, verbal stimuli can be processed as distinct, memorable units. For example, when the subjects in Whittlesea's Experiment 6 (Whittlesea, 1987; Whittlesea & Brooks, 1988; Whittlesea & Cantwell, 1987) were encouraged to treat pseudowords as whole units, a later test of the visual perception of new pseudowords showed striking similarity-to-old effects. When the initial experience with the pseudowords involved comparing them one letter at a time with a pseudoword that was prototypical for the training set, the perception of new pseudowords was facilitated in proportion to their similarity to the prototype, rather than their similarity to the training instances. In our current experiments, the feature list stimuli were also presented in such a way as to break up the lists into separate features and minimize the amount of item-distinctive processing. The relevant features were always in the same place in the list, no unique irrelevant features were included, and no memory for unique information was required, as had been done in Experiment 1. If we wanted to produce similarity effects with the feature list material, we presumably would have to change at least some of these processing characteristics.

Although it is at least useful to provide perceptually coherent whole items, it is probably important that there is not a predictive overall similarity among all members of the category. That is, the relevant features should not be perceptually integratable independently of the irrelevant features, a stricture very much in the spirit of Garner's (1976) work on the sorting of integral and separable dimensions. Because the relevant features of the current material do not seem to integrate into something resembling an overall shape, one has to keep checking the features individually. This seems to be characteristic of many medical categories; the relevant, verbally specified features look different in combination with different irrelevant, or at least, unspecified variables.

Taken together, the conditions that seem conducive to similarity effects are that (a) testing take place in a context of largely familiar items, (b) the rule be as complex as is often used for initial instruction, (c) the items be easily integrated into mnemonic units, (d) the relevant features not form a coherent perceptual whole independent of the irrelevant features, and (e) that extended experience possibly include a bias toward repetition of old items. We believe that, together, these conditions define a domain of considerable interest.

### *Prior Processing Episodes*

We have described the effects of previous instances as occurring because the items presented to our subjects acted as retrieval cues for prior processing episodes involving those instances, as well as gave information for applying a rule. Using the term "prior processing episodes," in contrast with the term "instance" is meant to stress the processing-dependent nature of both similarity and retrieval. What makes a prior instance similar and available in the current context is not its objective similarity, but the extent to which these items have been processed the same way. This approach is similar to research on episodic memory tasks, which is characterized by very close attention to specifically what was done with a stimulus on a prior trial, the exact conditions of retrieval, and the way in which these two fit with one another. This approach is commonly referred to as encoding or processing specificity (Kolers, 1979; Tulving & Thompson, 1973). Such work demonstrates the importance of micro-context—small variations in the setting and processing context of an item.

A processing-episode approach has proven useful in a variety of areas that do not directly require memory for particular episodes. For example, several of the experiments in this paper were modeled after experiments by Jacoby and his colleagues, on such topics as word identification (e.g., Jacoby 1983a, 1983b), judgments of fame (Jacoby, Kelley, Brown, & Jasechko, 1989), and judgments of the loudness of background noise (Jacoby, Allan, Collins, & Larwill, 1988). Whittlesea has shown processing-dependent episodic effects in perception of pseudowords, and in the word-superiority effect (Whittlesea & Brooks, 1988; Whittlesea & Cantwell, 1987). Perception of line drawings (Jacoby, Baker, & Brooks, 1989) has also shown processing-dependent retrieval effects. Jacoby and Brooks (1984) provide a general review, and Brooks (1987) offers a more specific application to categorization tasks. Overall, this work concurs with the findings of the current paper in suggesting that episodic effects are a prominent part of many tasks in which they are not definitionally required.

The distinction between "prior processing episodes" and "instances" is not strongly tested in this paper. We have demonstrated that the familiarity of the overall test list is important in controlling the size of episodic effects. But to have provided a strong test, we would have had to vary the processing of particular prior items, and we would have had to show that the same prior item had more of an effect when those conditions of processing were matched at the moment of test. Malt (1989) devised a very useful priming paradigm to investigate prototype versus exemplar strategies on-line. She found that old items that had been primed by presentation of a new similar item were classified faster than unprimed old items. However, this priming did not occur if the question asked of the immediately preceding new similar item was different from the regular classification question ("big or small?" rather than "A or B?"). Thus, more of the characteristics of the prior processing episode than simply the literal stimulus were important in affecting classification judgments.

We have adopted the "processing episode" terminology because the approach has proven useful in related work. It

also serves as a way of stressing that we are not assuming literal, "positivist" similarity when we refer to similarity effects.

### *Relation to Other Hybrid Models*

This work is hardly unique in suggesting that the operative form of knowledge is some form of hybrid between rule or prototype and instance or episodic knowledge (e.g., Medin, Dewey, & Murphy, 1983). We see our work as part of an ongoing effort to embed this suggestion into ecological and process rationales. From some of our own efforts, we can testify that investigating the hybrid question with the standard concept-learning materials and procedures does not lead to satisfying parallels to everyday problems. The episodic rationale just outlined, and the ecological rationale of previous sections are one attempt to generate more satisfactory preparations. However, other approaches in the current literature have made interesting progress on the same general problem.

One approach to instantiating the hybrid suggestion is Ross's work on reminders in problem solving (Ross, 1984, 1987, 1989; Ross & Kennedy, 1990). His work has shown that performance while learning a word-processing program or learning to solve simple probability problems is influenced by similarity to previously encountered problems in the same domain. Ross (1987) demonstrated that, while superficial similarity has an effect (e.g., whether the story line for the current and similar problem is the same), there was independent analytic competence (e.g., similarity of story lines had no effect if the correct formula was provided with the test problem), as well as a more analytic effect of similarity not tested in the current studies (the similar problem influenced the identification of the objects in the story with variables in the formulae). The parallel of his work with the current results is interesting because of the very different level of challenge being provided for the learner. In Ross's work, understanding the problem and the algorithms is the issue. For those problems there is doubt as to whether the person will retrieve the correct formula and will identify the correct semantic objects with the variables in the formula. The purpose of his research is to elucidate the processes by which the learner accomplishes these tasks, and similarity to specific prior problems is clearly implicated. In the current task, the analytic competence of the learner is hardly the issue. There is only one rule, it is an easy rule, and the identification of the item features with the elements of the rule is not in doubt. Rather, the issue here is to make its application faster or also to specialize it around previously experienced exemplars. It would seem that within-domain similarity is important for both levels of problem difficulty. Thus, we regard our work as complementary to that of Ross. Both demonstrate a role for prior instances, but they have different rationales and apply to different circumstances.

In another approach, Logan (1988) published a series of experiments that address instance effects in automaticity. He used tasks like basic arithmetic, in which he saw two ways to get an answer: (a) the purposeful use of an algorithm and (b)

automatic, instance-based memory for one's previous answer to the same question. Performance is based on a race between the two processes. As practice continues, the base of instances increases, thereby increasing the probability that the fastest access of an instance is fast enough to beat the algorithm. Automatic performance is the extreme case of this race process, where there are so many episodes in the knowledge base that the instance-based process virtually always wins the race. Thus, he equates automatic performance with instance-based processes and purposeful performance with algorithm-based processes. This race between prior instances and a known algorithm is obviously similar to the situation in our experiments. The simultaneous effects of rules and prior episodes could be looked at either as a very early version of the process he describes or as a stretching of the word "automaticity." We tend toward the former view, particularly because the lack of awareness of the source of errors on the negative-match items suggests that they were not produced by a deliberate strategy.

Although Logan's views seem fundamentally compatible with our experiments, there are some complementary differences between his research and ours. Conceptually, his work describes a pure instance model that has none of the coding specificity or retrieval variations discussed in the previous section. Most important for method is that he treats instances as statistical entities; he does not attempt to keep track of the influence of individual items. While this does not allow the specificity of effects provided by the negative matches in our studies, he is also not tied to a small number of specific items. Although he has not yet directly addressed the problem of transfer to new instances, the manner in which he could extend his explanation seems clear and could provide interesting convergence with the results of our methodological approach.

### Future Directions

The specific materials used in this experiment were designed to favor rule application. While this has value in eliminating difficulty of rule application as a necessary condition for similarity effects, a minimal additive rule and unambiguous values on binary dimensions are not the conditions of ultimate interest. As a first step in extension, we have already completed parallel studies teaching dermatological classifications to medical students (Brooks, Norman, & Allen, 1990). These studies have also shown an effect of similar prior instances on the classification of new photos of dermatological lesions and possibly on the description of ambiguous features within the new photos. This is relevant to the current article in that it indicates that the current results are not limited to simple, artificial materials.

A potentially interesting extension would require a considerably more complex and varied rule set, for which the dermatological material might also serve. With as simple a rule as was used in the current article, we have had to content ourselves with the claim that the operation of rules specializes around simultaneously evoked exemplars. But with a more complicated rule structure, the operation of episodic retrieval might produce specialization of the rule itself. In particular

processing episodes, only some portion of a complicated rule would be used. Conditions in the rule that had not been evoked with stimuli similar to the current probe, but which would then become relevant might not easily be available. Some segments of the original rule "just wouldn't come to mind," producing what would be experienced as "silly mistakes." With enough experience in a context that did not produce a lot of such errors, the original complete rule might be retrieved only with difficulty. In such cases of selective evocation, we could claim that the rule itself had specialized into smaller, less general, operational rules.

The same process of supplementing the original rule guidance with a dose of episodic information should be even more important when the rule is only a heuristic that is clearly wrong in some cases. If the exceptional cases are the majority, and the demand characteristics of the experiment suggest the need to analyze, we would undoubtedly see hypothesis testing and modification as the learner tried to improve the rule. But, as was approximated in Experiment 4, if the learner is not directly challenged, the exceptions not too bizarre, and most of the items are familiar "friends," then accepting an occasional boost from a prior similar episode might pass nearly unanalyzed. Adapting and refining a rule in a neighborhood of friends is likely to be a much less analytic affair.

In the experiments in this article, we seem to have a situation in which people are not particularly reflective. They do not spontaneously change their behavior to eliminate errors to negative matches, and they do not even seem to recognize negative matches as the source of their errors. The threat of making errors to real-world versions of negative matches is a major reason for forcing oneself to apply a rule, to not just rely on immediate appearance. In this case, the subjects were not making enough errors to induce them to think about the source of the errors they did make, and in any event many of these errors were to old items. When we did direct their attention to the errors on the negative matches (Experiment 3), the rule seemed to provide a source of attributions that pointed away from overall similarity to whole items ("I guess the long neck made the animal tall, as if it had long legs," as stated by one subject). The conditions of this experiment seem more like the nonreflectiveness of everyday life, as opposed to the reflective, problem-solving behavior that we often elicit from our subjects in concept-learning experiments.

### References

- Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Broadbent, D. E. (1989). Lasting representations and temporary processes. In H. L. Roediger & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 211-223). Hillsdale, NJ: Erlbaum.
- Broadbent, D. E., FitzGerald, P., & Broadbent, M. H. P. (1986). Implicit and explicit knowledge in the control of complex systems. *British Journal of Psychology*, 77, 33-50.
- Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 169-211). Hillsdale, NJ: Erlbaum.

- Brooks, L. R. (1987). Decentralized control of categorization: The role of prior processing episodes. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 141-174). Cambridge, England: Cambridge University Press.
- Brooks, L. R., Norman, G. R., & Allen, S. W. (1990). *The role of specific similarity in a medical diagnostic task*. Manuscript submitted for publication.
- Garner, W. R. (1976). Interaction of stimulus dimensions in concept and choice processes. *Cognitive Psychology*, 8, 98-123.
- Hartley, J., & Homa, D. (1981). Abstraction of stylistic concepts. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 33-46.
- Homa, D. (1978). Abstraction of ill-defined form. *Journal of Experimental Psychology: Human Learning and Memory*, 4, 407-416.
- Homa, D., Sterling, S., & Trepel, L. (1981). Limitations of exemplar-based generalization and the abstraction of category information. *Journal of Experimental Psychology: Human Learning and Memory*, 7, 418-439.
- Jacoby, L. L. (1983a). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior*, 22, 485-508.
- Jacoby, L. L. (1983b). Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 21-38.
- Jacoby, L. L., Allan, L. G., Collins, J. C., & Larwill, L. K. (1988). Memory influences subjective experience: Noise judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 14, 240-247.
- Jacoby, L. L., Baker, J. G., & Brooks, L. R. (1989). Episodic effects on picture identification: Implications for theories of concept learning and theories of memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 275-281.
- Jacoby, L. L., & Brooks, L. R. (1984). Nonalytic cognition: Memory, perception and concept learning. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory: Vol. 18*. (pp. 1-46). San Diego, CA: Academic Press.
- Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Why memory is misleading. *Journal of Personality and Social Psychology*, 56, 326-338.
- Kolers, P. A. (1979). A pattern-analyzing basis for recognition memory. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing and human memory* (pp. 363-384). Hillsdale, NJ: Erlbaum.
- Lewicki, P. (1986). Processing information about covariations that cannot be articulated. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12, 135-146.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95, 492-527.
- Malt, B. C. (1989). An on-line investigation of prototype and exemplar strategies in classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 539-555.
- Medin, D. L., Dewey, G. I., & Murphy, T. D. (1983). Relationships between item and category learning: Evidence that abstraction is not automatic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 607-625.
- Nosofsky, R. M., Clark, S. E., & Shin, H. J. (1989). Rules and exemplars in categorization, identification, and recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 15, 282-304.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 219-235.
- Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum.
- Ross, B. H. (1984). Reminders and their effects in learning a cognitive skill. *Cognitive Psychology*, 16, 371-416.
- Ross, B. H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 629-639.
- Ross, B. H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 456-468.
- Ross, B. H., & Kennedy, P. T. (1990). Generalizing from the use of earlier examples in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 42-55.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127-190.
- Tulving, E., & Thompson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352-373.
- Whittlesea, B. W. A. (1987). Preservation of specific experiences in the representation of general knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13, 3-17.
- Whittlesea, B. W. A., & Brooks, L. R. (1988). Critical influence of particular experiences in the perception of letters, words, and phrases. *Memory and Cognition*, 16, 387-399.
- Whittlesea, B. W. A., & Cantwell, A. L. (1987). Enduring influence of the purpose of experiences: Encoding-retrieval interactions in word and pseudoword perception. *Memory and Cognition*, 15, 465-472.



## Appendix

## Logical Descriptions of Experimental Stimuli

Item number	Body shape	Spots	Leg length	Neck length	Number of legs	Background
1	1	1	1	0	0	4
2	0	1	1	0	1	1
3	1	0	1	1	1	2
4	1	1	0	1	0	3
5	1	0	0	0	1	2
6	0	1	0	1	1	1
7	0	0	1	1	0	4
8	0	0	0	0	0	3
9	1	1	1	1	1	2
10	1	0	1	0	0	4
11	1	1	0	0	1	2
12	0	1	1	1	0	4
13	0	1	0	0	0	3
14	0	0	1	0	1	1
15	0	0	0	1	1	1
16	1	0	0	1	0	3

*Note.* The 1s and 0s represent the two different values on the dimensions. Body shape, spots, and leg length are the dimensions that are relevant with respect to the rule, while neck length and number of legs are irrelevant. Notice that items 1 through 8 have matching items among numbers 9 through 16, identical to all dimensions except spots. All other items differ by at least two dimensions. The assignment of items to conditions for Experiments 1, 3, and 4 was as follows (in Experiment 2, only Rule 1 was used):

Rule 1: positive old—1, 3, 6, 8;      positive match—10, 9, 15, 13;  
           negative old—2, 4, 5, 7;      negative match—14, 16, 11, 12.

Rule 2: positive old—11, 12, 14, 16;  
           negative old—9, 10, 13, 15;

Rule 3: positive old—9, 10, 13, 15;      positive match—3, 1, 8, 6;  
           negative old—11, 12, 14, 16;      negative match—5, 7, 2, 4.

Rule 4: positive old—2, 4, 5, 7;      positive match—14, 16, 11, 12;  
           negative old—1, 3, 6, 8;      negative match—10, 9, 15, 13.

Received January 18, 1989  
 Revision received April 2, 1990  
 Accepted April 19, 1990 ■