Recognition of Exceptions and Rule-Consistent Items in the Function Learning Domain

Edward L. DeLosh (delosh@lamar.colostate.edu)
Department of Psychology, Colorado State University
Fort Collins, CO 80523 USA

Abstract

Recent studies suggest that participants commonly abstract rules when learning concepts, but a remaining question is whether they retain and apply knowledge of individual instances subsequent to rule abstraction. Research in the category learning domain indicates that exemplar information is retained and that exceptions to a category rule have special status in memory (Palmeri & Nosofsky, 1995). The present experiment examines whether these findings extend to function learning. Participants learned associations between stimulus and response magnitudes that were related according to a negative linear function. Twelve stimulus-response pairs were given, some consistent with the negative linear rule, others exceptions to the rule. After each of six training sessions, previously studied stimulus magnitudes were presented as tests of learning accuracy. Participants were also given extrapolation trials followed by a final recognition test that included old and new rule-congruent and rule-incongruent items. Extrapolation was extensive. In addition, analyses revealed poorer learning and recognition for exceptions than for rule-congruent items, plus a high rate of false alarms for new rule-congruent items. These findings suggest that although the conceptual knowledge acquired in function learning tasks centers on rules, exceptions to these rules do not have special status in memory.

Introduction

In the spirit of classic hypothesis-testing models of classification learning (e.g., Bower & Trabasso, 1963; Levine, 1975; Restle, 1962), contemporary theories have revitalized the idea that conceptual behavior is based on the abstraction and application of rules. Recent rule-based models developed by Nosofsky, Palmeri, and McKinley (1994) and DeLosh, Busemeyer, and McDaniel (1997) have been successful in accounting for a variety of data in the category and function learning domains, respectively. Both models propose that conceptual behavior reflects the joint influence of exemplar and rule-based processes. This emerging theoretical approach begs the following empirical questions: In what way do rules and exemplars jointly contribute to conceptual behavior? Are individual instances learned and retrieved subsequent to rule abstraction? Is conceptual behavior characterized by individual differences in the use of rules versus exemplars? The present experiment considers these issues as they apply to the function learning domain.

Rule Abstraction in Function Learning

Functions are abstract concepts that characterize the relationship between two causal variables. A function maps a set of input values on a stimulus continuum into a set of output values on a response continuum such that each input value is assigned only one output value. In a typical function learning task, input and output dimensions are related according to a simple mathematical function. Learning occurs on a trial-by-trial basis through experience with individual input-output pairs. In DeLosh et al. (1997), for example, participants learned associations between drug dosages and the magnitude of clinical effect caused by those dosages. The dosage-effect relationship was either linear, exponential, or quadratic. On each learning trial, a drug dosage was represented on a computer monitor as a bar length. Participants then predicted the magnitude of effect for that dosage by changing the length of a second bar. Then they were shown the "correct" magnitude of effect (represented by a third bar) as defined by the objective function. Numerous trials of this type were given, such that each of many dosage-effect pairs was presented several times.

Learning in this type of task potentially involves memory for specific input-output pairs, abstraction of relational information pertaining to the input and output dimensions, or some combination of these processes. DeLosh et al. (1997) examined these possibilities by presenting a series of extrapolation tests. Participants were given new dosage values outside the range of those given during learning. They responded to these extrapolation stimuli by generating outputs beyond the range of learned responses, and did so in a manner consistent with the form of the assigned function. A pure exemplar-based model of function learning (i.e., an extension of ALCOVE; cf. Kruschke, 1992) was unable to extrapolate to the extent observed with participants, revealing the necessity for rule learning (i.e., the abstraction of relational information during acquisition) or rule-based responding (i.e., the abstraction of relational information during retrieval) instead of or in addition to exemplar learning.

In a second line of research, DeLosh (1994) observed discontinuous patterns of responding during function learning, and these discontinuities were similar for a condition with explicit hypothesis-testing instructions and a condition with standard free-strategy instructions. This observation lends support to the idea that function learning involves the systematic sampling and testing of global input-output rules. It appears, then, that rule abstraction plays a central role in the learning and application of function-based concepts. But to what extent is exemplar information retained and used subsequent to rule abstraction?

RULEX Model of Category Learning

This issue has recently been examined as it applies to category learning. Nosofsky et al. (1994) proposed a rule-based model of classification (RULEX) in which participants abstract simple logical rules based on single dimensions or conjunctions of dimensions, supplemented by memory for exceptions to those rules. Because members of ill-defined categories can not be classified based solely on the application of logical rules, information pertaining to exceptions is central to the success of the model. The model therefore assumes that there is residual memory for old exemplars and that old exceptions have special status in memory. Consistent with this assumption, Palmeri and Nosofsky (1994) observed intact memory for old exemplars, and better recognition memory for old exceptions than for old rule-congruent items.

Note, however, that there are several differences between the category learning tasks examined in the above studies and the function learning task considered in the present experiment. In *category learning*, responses consist of discrete and nominal categories that do not have any numerical status. Rules learned in a category learning task are logical rules for mapping stimuli onto arbitrary response categories (e.g., red and square stimuli belong to Category A). In *function learning*, responses lie on a continuum and are numerically related to one another and to stimuli. Therefore the rules abstracted in a function learning task may reflect the numerical relationship between stimuli and responses (e.g., drug dosage is positively correlated with heart rate).

Despite these differences, a plausible explanation of function learning is that participants abstract a functional rule and memorize exceptions to that rule, comparable to the processing assumptions of RULEX. One might ask, then, do the findings of Palmeri and Nosofsky (1994) generalize to function learning? Is exemplar information retained subsequent to rule abstraction? Do exceptions to function-based rules have special status in memory?

Overview of the Current Experiment

The current experiment examines these questions by extending on the method used by DeLosh et al. (1997). Participants learned associations between stimulus and response magnitudes (i.e., bar heights) that were related according to a negative linear function. Twelve stimulus-response pairs were given, some consistent with the negative linear rule, others exceptions to the rule. After each of six training sessions, previously studied stimulus magnitudes were presented as tests of learning accuracy. During the final test session, participants were also given extrapolation trials to test for rule abstraction. To examine memory for individual stimulus-response pairs, participants were then given a final recognition test that consisted of old rule-congruent items, old exceptions, new rule-congruent items, and new rule-incongruent items. A random-mapping condition was included to examine performance when learning is strictly based on memory for individual input-output pairs.

Method

Participants and Apparatus

Sixty-eight Colorado State University undergraduates participated in partial fulfillment of a requirement for an introductory psychology course. Participants were tested in pairs or groups of three in a laboratory room equipped with three computer workstations. Stimuli were presented on a 14" color monitor at a distance of approximately 60 cm and responses were collected using a standard computer keyboard placed on the desk in front of the monitor. A computer program controlled the presentation of the instructions and stimuli as well as the collection of participants' responses.

Design

The experiment included two conditions based on the mapping between stimulus and response magnitudes. In a functional-mapping condition, stimulus and response magnitudes were related according to the negative linear function $y = 200 - 1.7x$. The random-mapping condition included the same stimulus and response magnitudes used in the functional-mapping condition, but stimuli were randomly paired with responses. Mapping was manipulated between participants, with 35 participants randomly assigned to the functional-mapping condition and 33 to the random-mapping condition.

Stimuli and Responses

All stimuli and responses were presented in the form of vertical bars with the height of each bar proportional to the assigned stimulus or response magnitude. The range of possible stimulus magnitudes was 0 to 100 as indicated by an unfilled vertical bar labeled 0 to 100. The range of possible responses was 0 to 200 as indicated by an unfilled vertical bar labeled 0 to 200. Due to limited resolution of the computer monitor, response magnitudes were constrained to integer values.

A total of 12 stimulus-response pairs were given during training. For the functional-mapping condition, 10 of the stimulus-response pairs corresponded to the negative linear rule. Two were exceptions to the rule (Pairs 4 and 9; see Figure 1), with the learned response deviating from the rule-defined response by 42 units. For the random-mapping condition, the stimulus and response magnitudes used for rule-congruent items in the functional-mapping condition were randomly paired. Two random mapping sets were generated: 16 participants received Set 1 and 17 received Set 2. Pairs 4 and 9 of the random-mapping sets were identical to the exceptions used the functional-mapping condition. Figure 1 provides a graphical representation of the stimulus-response pairs used in the functional-mapping set and one of the random-mapping sets.

Note that the specific stimulus magnitudes given during training ranged from 22.5 to 77.5 on the 0 to 100 stimulus scale, with corresponding response magnitudes ranging from 68 to 162 on the 0 to 200 response scale. Values beyond this range were given after the final learning session to examine extrapolation. The following stimulus magnitudes were used on extrapolation trials: 2.5, 7.5, 12.5, 17.5, 82.5, 87.5, 92.5, and 97.5.

Procedure

At the beginning of the experiment, participants read a set of instructions on the computer monitor. In the instructions, participants were told that they would observe a hypothetical pharmacology experiment in which dosages of an

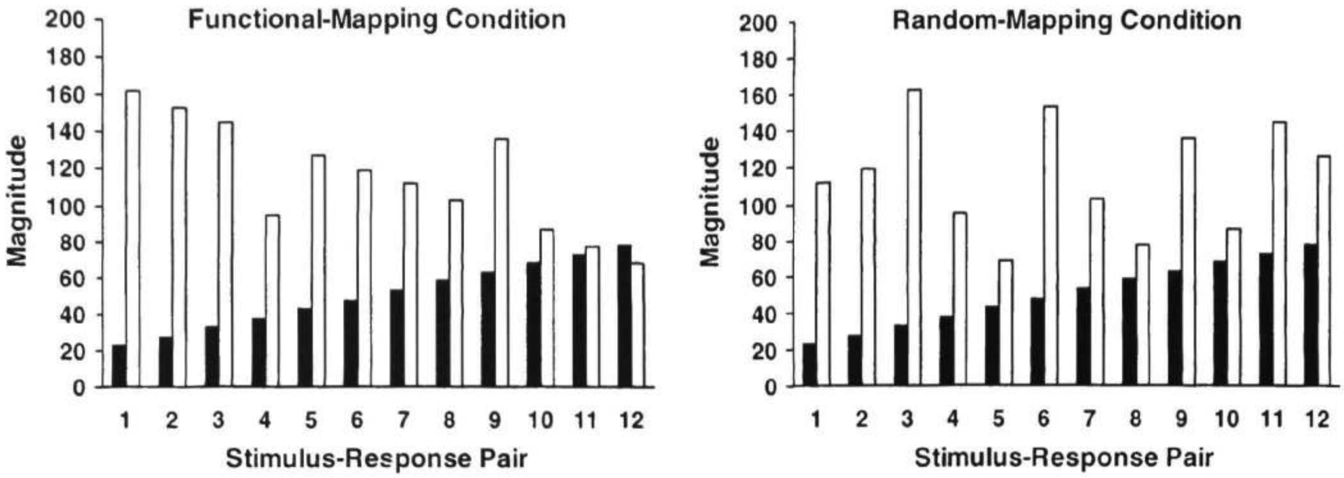

Figure 1: Graphical representation of the stimulus-response pairs given in the functional- and random-mapping conditions. Pairs 4 and 9 are exceptions.

unknown drug are given to subjects and the level of arousal produced by each dosage is measured. They were instructed to predict the level of arousal produced by each drug dosage given, and when given feedback, to remember the level of arousal associated with each dosage. They were never told to figure out the relationship between dosage and arousal levels, or that a systematic relationship might exist. With regard to the learning task itself, the instructions described the format of the presentation screen and the appropriate keys for making a prediction. Once these instructions were understood, a sample trial was given in order to familiarize the participant with the presentation screen and response procedure.

After the sample trial, participants proceeded through alternating training and test sessions, with each of six training sessions followed by a test session. During training, the stimulus magnitudes that constituted the stimulus set were presented one at a time in random order. For each trial, three unfilled vertical bars were presented simultaneously on the monitor. The leftmost bar was titled Drug Dosage and had tick marks and value labels every twenty units from 0 to 100, and the remaining two bars were titled Predicted Level of Arousal and Observed Level of Arousal, respectively, with tick marks and value labels every twenty units from 0 to 200. The relative lengths of these unfilled bars on the screen were proportional to the number of units they represented.

On a given trial, the left bar was filled in from the zero point (at the bottom of the bar) to the input value representing the amount of drug administered. Participants then used the arrow keys on their keyboard to fill in the second vertical bar from the zero point to the desired prediction value, and pressed the space bar when finished. Participants were allowed as much time as needed to make their prediction. Once the space bar was pressed, the correct level of arousal (i.e., the response value assigned to the stimulus according to the mapping condition) was shown on the rightmost vertical bar, along with an accuracy score of 0 to 100, computed as 100 minus the square of the participant's prediction error. This correct-response feedback was displayed for 6 s. The next trial was initiated by pressing the enter key.

After participants completed a training session, they were given 12 test trials consisting of the exact stimulus values shown during training. The test stimuli were presented in random order. Individual trials proceeded in exactly the same fashion as training trials, except the rightmost bar (for presenting the correct response magnitude) was not included and no other feedback was provided. During the sixth and final test session, a sequence of eight extrapolation trials was given after the standard test trials. These extrapolation trials were given in random order, and like test trials, feedback was not provided.

The experiment concluded with a final yes-no recognition test. A series of 24 stimulus-response pairs was shown on the computer monitor. For each item participants were instructed to respond yes (press the "y" key) if they believed the pair was previously given during the experiment or no (press the "n" key) if they believed the pair was not given during the experiment. The recognition test consisted of 12 old items (each shown a total of 12 times during the alternating training and test sessions) and 12 new items. For the functional-mapping condition, these items can be grouped into four types: *old rule-congruent items* (the 10 rule-generated pairs given during training), *old rule-incongruent items* (the 2 exceptions given during training), *new rule-congruent items* (2 lures with the same stimulus values as exceptions, but paired with the appropriate rule-generated response), and *new rule-incongruent items* (10 completely new pairs inconsistent with the negative linear rule). The 24 recognition trials were presented in random order.

Results

Learning

In order to compare learning performance for exceptions versus non-exceptions, the average absolute prediction error on test trials was computed as a function of item type and test session for each participant. These averages were submitted to a 2 x 2 x 6 (Mapping x Item Type x Test Session) mixed analysis of variance (ANOVA). The rejection level was set at .05 for this and all other analyses reported in the current study. Main effects of mapping [$F(1,66) = 16.49$, $MSE = 428.41$], item type [$F(1,66) = 10.87$, $MSE = 257.86$], and test session [$F(5,330) = 15.66$, $MSE = 150.89$] were obtained. As observed in previous experiments (e.g., Carroll, 1963; DeLosh, 1996), performance was better in the functional-mapping condition than in the random-mapping condition, and improved from test session to test session. In addition, prediction accuracy was better for non-exceptions than for exceptions.

A significant interaction between mapping and item type [$F(1,66) = 70.44$, $MSE = 150.89$] was also observed. In the random-mapping condition, performance was better for exceptions (*Mean Prediction Error* = 24.85) than for randomly paired stimulus-response values ($M = 30.58$). In the functional-mapping condition, performance was better for rule-congruent items ($M = 15.25$) than for exceptions ($M = 28.40$). Therefore, the specific stimulus-response pairings used as exceptions were easier to learn than random pairings, but despite this, participants were less accurate with exceptions than with rule-congruent items in the functional-mapping condition (see Figure 2; also see Figure 3).

The analysis also yielded a significant interaction between mapping and test session [$F(5,330) = 2.36$, $MSE = 150.89$], revealing greater improvement over test sessions for the functional-mapping condition than for the random-mapping condition. None of the remaining interactions were statistically reliable ($ps > .10$).

Recognition

In order to analyze recognition performance, hit and false alarm rates were computed for each participant in the functional- and random-mapping conditions. One participant in the functional-mapping condition terminated the experiment prior to completing the recognition test, therefore the following analyses are based on the remaining 67 participants. Independent-sample t tests revealed a higher hit rate for items in the functional-mapping condition ($M = .80$) than for items in the random-mapping condition ($M = .73$). The false alarm rate did not significantly differ across conditions ($p > .10$; $M = .24$ and .30, respectively).

A more detailed analysis of the recognition data was then conducted for the functional-mapping condition. Hit and false alarm rates were computed for the four types of items given on the recognition test, yielding the means given in Table 1. The hit rate for old rule-congruent items significantly differed from that of old rule-incongruent items [$t(33) = 4.29$], such that recognition memory for rule-based items

Table 1: Hit and false alarm rates for each type of item in the functional-mapping condition.

Item type	Proportion of yes responses
Old rule-congruent items	.89
Old rule-incongruent items	.59
New rule-congruent items	.59
New rule-incongruent items	.17

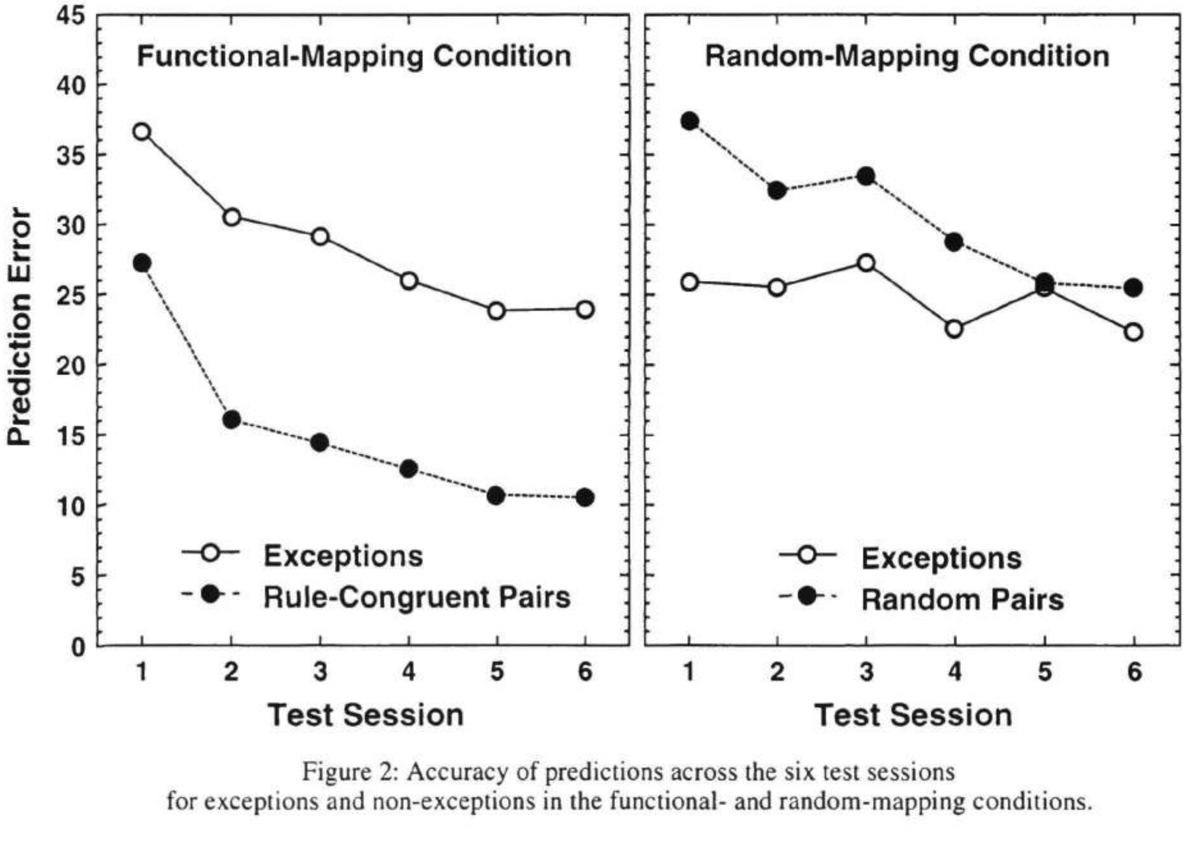

Figure 2: Accuracy of predictions across the six test sessions for exceptions and non-exceptions in the functional- and random-mapping conditions.

was better than recognition memory for exceptions.[1] In addition, the false alarm rate differed for new rule-congruent items and new rule-incongruent items, $t(33) = 7.00$. Items consistent with the negative linear rule produced a higher rate of false recognition than was observed with other new items. In fact, the false alarm rate for the rule-congruent lures did not differ from the hit rate for exceptions ($p > .10$).

Extrapolation

Performance on extrapolation trials was examined to further assess whether learning in the functional-mapping condition was based on rule abstraction. One participant did not complete the extrapolation trials, so the following data are based on the remaining 34 participants. Figure 3 shows the average of participants' predictions as compared to the objective responses across all stimulus values from the last test session. The stimulus magnitudes within the dotted lines correspond to stimuli given during learning; those outside the dotted lines are extrapolation stimuli.

The figure reveals that participants extrapolated well beyond the range of learned responses in both extrapolation regions, extending 16 and 24 units beyond learned responses in the low and high extrapolation regions, respectively. Moreover, the extrapolation responses were highly systematic, closely approximating the objective function. Participants' predictions were, in fact, closer to the objective values for extrapolation stimuli ($M = 12.19$) than for exceptions ($M = 19.49$). This extensive extrapolation replicates past findings (DeLosh, 1994, 1996) and provides strong evidence for rule abstraction (see Discussion).

It is also noteworthy that extrapolation responses in the random-mapping condition were positively correlated with stimulus magnitudes (M = 43.85, 62.82, 83.12, 79.33, 106.42, 109.94, 114.94, and 117.42 for extrapolation trials 1 through 8, respectively). This suggests that participants attempted to apply a rule even in the random-mapping condition. Moreover, the particular pattern of extrapolation is consistent with findings in the function-learning literature that reveal biases toward increasing monotonic functions (Brehmer, 1974; Busemeyer et al., 1997).

Individual Differences

To determine whether the average data described above is representative of individual learners, extrapolation performance was examined for each of the 34 participants in the functional-mapping condition. Five of these participants deviated from the group data, failing to extrapolate in the two extrapolation regions. Contrary to the large advantage for rule-congruent items observed in the group data, these participants also showed: (a) similar learning performance for rule-congruent items (*Mean Prediction Error* = 20.95) and exceptions ($M = 26.57$); (b) a similar hit rate for old rule-congruent items ($M = .84$) and exceptions ($M = .80$); and (c) a similar false alarm rate for new rule-congruent ($M = .40$) and rule-incongruent items ($M = .34$).

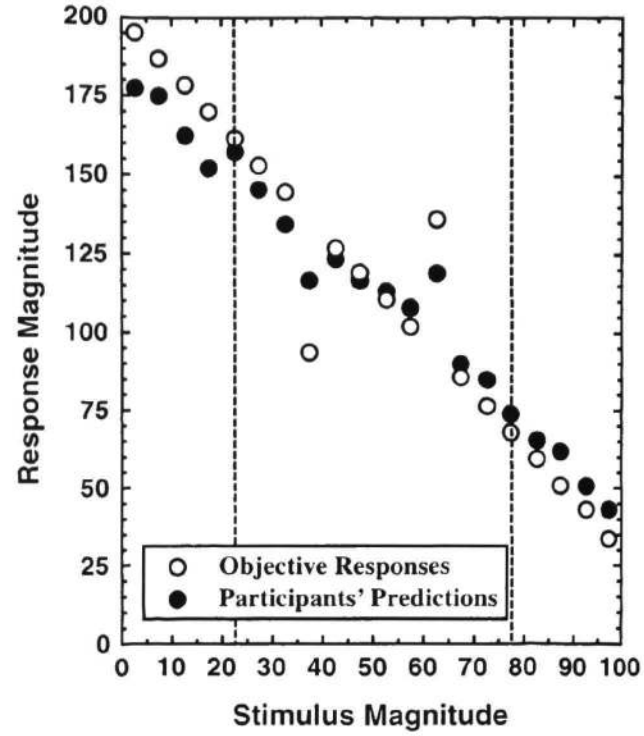

Figure 3: Participants' final test predictions for the functional-mapping condition.

Discussion

Rule Abstraction in Function Learning

The present findings support the view that participants often abstract and apply rules when learning function-based concepts, at least for simple functional relations. First consider the extrapolation results. The observed pattern of extrapolation responses approximates the objective negative linear function and therefore suggests that participants abstracted and applied information about the stimulus-response relationship. In support of this interpretation, DeLosh et al. (1997) formally tested a pure exemplar-based model of function learning (an extension of ALCOVE; see Kruschke, 1992) and showed that the model can not account for extensive extrapolation, as observed here. In order to produce extensive extrapolation, it is necessary to include a rule-based mechanism in which relational information is abstracted during the learning or retrieval of individual instances (cf. Busemeyer et al., 1997; DeLosh et al., 1997).

The present experiment also provides new corroborative support for rule abstraction. If participants only learn and remember individual stimulus-response pairs, one might not expect differences between the functional- and random-mapping conditions. There was, however, an advantage for the functional-mapping condition in both learning and recognition. Similarly, if participants rely on exemplar learning even in the functional-mapping condition, one might not expect differences between rule-congruent items and exceptions. In the random-mapping condition in which participants were required to memorize individual instances, there was no advantage for rule-congruent items. However, in the

---

[1] For comparison, there was no difference between the hit rate for exception pairs ($M = .75$) and random pairs ($M = .70$) in the random-mapping condition ($p > .10$).

functional-mapping condition there was an advantage for rule-congruent items over exceptions in both learning and recognition. In addition, the experiment yielded a high rate of false alarms for new rule-congruent items. Participants often judged rule-congruent lures as having occurred before, and did so at a rate equivalent to that of exceptions that were shown 12 times during acquisition.

Note, however, that a few learners deviated from the group averages described above. These participants did not extrapolate beyond the range of learned responses in the two extrapolation regions. As discussed by DeLosh et al. (1997), this failure to extrapolate may reflect a strategy that centers on exemplar learning (also see DeLosh 1994, 1996). In any case, responding does not appear to be based on rules for this subset of participants. If these participants do not abstract and use rules, one would expect similar learning and recognition performance for rule-congruent and rule-incongruent items (in contrast to the large advantage for rule-congruent items observed in the group data). This is precisely what was found. It appears, then, that the large majority of participants abstract and apply rules, but a few may rely exclusively on memory for individual instances.

Memory for Instances in Function Learning

Although rule abstraction appears to play a central role in function learning for most participants, results show that these participants also have residual memory for individual instances. Within the functional-mapping condition, old rule-congruent items were more likely to be judged as having occurred before than were new rule-congruent items. Likewise, old rule-incongruent items (i.e., exceptions) were more likely to be judged as having occurred before than were new rule-incongruent items. This indicates that participants do not simply make recognition decisions by judging whether an item is consistent or inconsistent with the abstracted rule. Rather, recognition judgments appear to be based, at least in part, on familiarity with or recollection of individual stimulus-response pairs.

Memory for Exceptions to Function-Based Rules

Even though participants in function-learning tasks do appear to retain exemplar information, exceptions do not seem to have special status in memory. Contrary to findings from category learning experiments (see Palmeri & Nosofsky, 1995), rule-congruent items were better learned and better recognized than exceptions. In fact, performance for stimulus-response pairs that were used as exceptions was much worse when those pairs were learned in conjunction with rule-generated pairs (the functional-mapping condition) than when learned in conjunction with random pairs (the random-mapping condition). It therefore appears that learning a function-based rule interferes with learning and memory for specific instances that are exceptions to that rule.

Conclusions

In sum, the current study supports the view that the learning and application of function-based concepts involves rule abstraction as well as memory for specific instances. One possible instantiation of this hybrid approach, following from the RULEX model of category learning, is that participants learn functional rules and remember exceptions to those rules. However, this particular rule-plus-exemplar account is not supported by the present experiment. Unlike findings from the category learning literature, exceptions to function-based rules do not appear to have special status in memory.

Acknowledgments

The author would like to thank Jayson Johns and Jason Lickel for their assistance in participant testing, data scoring, and preliminary data analysis.

References

Bower, G., & Trabasso, T. (1963). Reversals prior to solution in concept identification. *Journal of Experimental Psychology, 66*, 409-418.

Brehmer, B. (1974). Hypotheses about relations between scaled variables in the learning of probabilistic inference tasks. *Organizational Behavior and Human Performance, 11*, 1-27.

Busemeyer, J. R., Byun, E., DeLosh, E. L, & McDaniel, M. A. (1997). Learning functional relations based on experience with input-output pairs by humans and artificial neural networks. In K. Lamberts & D. Shanks (Eds.), *Knowledge, concepts, and categories*. East Sussex, UK: Psychology Press.

Carroll, J. D. (1963). *Functional learning: The learning of continuous functional maps relating stimulus and response continua* (ETS RB 63-6). Princeton, NJ: Educational Testing Service.

DeLosh, E. L. (1994). *Rule abstraction and hypothesis testing in the learning of functional concepts*. Master's thesis, Department of Psychology, Purdue University, West Lafayette.

DeLosh, E. L. (1996). *Effects of mnemonic variables on function and category learning*. Doctoral dissertation, Department of Psychology, Purdue University, West Lafayette.

DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 968-986.

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

Levine, M. (1975). *A Cognitive Theory of Learning: Research on Hypothesis Testing*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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

Palmeri, T. J., & Nosofsky, R. M. (1995). Recognition memory for exceptions to the category rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 548-568.

Restle, F. (1962). The selection of strategies in cue learning. *Psychological Review, 69*, 329-343.