When A’s and B’s are C’s and D’s: The effect of the cross-classification of items on learned concepts

Seth Chin-Parker (chinpark@s.psych.uiuc.edu)
Department of Psychology, 603 East Daniel Street
University of Illinois, Urbana-Champaign
Champaign, IL 61820 USA

Brian H. Ross (bross@s.psych.uiuc.edu)
Beckman Institute, 405 North Mathews Avenue
University of Illinois, Urbana-Champaign
Urbana, IL 61801 USA

Abstract

Category learning plays a crucial role in cognition because the acquired concepts are used later for categorization, explanation, communication, problem solving, and inference. However, there is little information about how category learning can affect previously acquired concepts. This issue is especially interesting when one considers items that belong to multiple categories, items that can be cross-classified. The current study investigates a learning situation where one classification set is learned and the knowledge gained is assessed. Then a second, orthogonal classification set is learned for the same items. The experiments show that there is an effect of the cross-classification on judgements made about items in terms of the initial classification set. There seem to be two possible effects of learning a cross-classification for items. The secondary learning causes either an intrusion of attribute information critical for the second classification set, or a specification as to what attributes are critical for the first classification set.

Introduction

Things we encounter in the world belong to multiple categories. The same person can be a runner, a mother, an intellectual. The same car can be a station wagon, a Toyota, a danger to pedestrians. This complex organization involving the cross-classification of items can be seen in common categories (Ross & Murphy, 1999), but its effect on category learning has been largely overlooked. In a typical category learning study, a subject learns to categorize a set of items into two categories. The focus has been on how the learning proceeds and what knowledge is acquired through the learning, as well as how the knowledge is represented by the learner. These are critical issues, but if the aim is to understand or model human category learning, more complex learning situations should be studied.

One type of complex learning situation that has received attention recently involves category learning in contexts where background knowledge has some bearing on the categories of interest (for reviews see Heit, 2001, and Murphy, 2002). However, the effect of knowledge in these studies is unidirectional. The background knowledge available to the learner has some effect on how category learning proceeds. The current study expands our understanding of how background knowledge interacts with current learning, focusing on whether current learning can affect established knowledge.

For example, a young child learns from both picture books and adult interaction which creatures in the world are “cats,” and which ones are “dogs.” The learning is extensive enough that the child can recognize easily what is a “cat” and what is a “dog,” even that a tiger is a categorized as a “cat” and a wolf as a “dog.” Later, the child learns that people keep some animals as “pets.” Now, the child learns which animals are pets and which are not. This category would cut across the “cat” and “dog” distinction that the child already has learned – some cats (and dogs) are pets while others are not. The question arises as to whether this later learning would have any predictable effect on category - based decisions the child makes about cats and dogs. The learning about “pets” and “not pets” might have no effect on how the child thinks about a cat she happens to encounter. However, learning about the critical attributes that determine “petness” might affect the child’s concept of cats (or dogs) in general.

An intelligent system uses past experience to make successful predictions about novel situations (Anderson, 1991), and category research has helped us understand how this takes place. At the same time, it is important to recognize that the learning that occurs in these new situations can interact with and possibly modify stored knowledge. An understanding of these interactions could help to specify how the system is representing knowledge and how best to model the processes at work. Studying the effects of cross-classification is one means of examining this issue.

Effects of Cross-Classification

Much study and thought has gone into understanding how people recognize the category membership of things. The focus of current category learning research has been largely on the learning of one classification set – an item encountered during the learning can belong to one of two contrasting categories. Although there is still debate on several key issues, it is generally agreed that when learning to classify items, subjects attend to diagnostic attributes, those attributes that are predictive of category membership (Tversky, 1977). These attributes are subsequently more salient in the category representation, having a greater effect on later classification decisions (Kruschke, 1992) and on typicality ratings of items (Chin-Parker & Ross, 2002).

The cross-classification of items adds an interesting and important dimension to these ideas. One possibility is that category learning proceeds independently for orthogonal classification sets. Given that a child already knows how to classify creatures as "cats" and "dogs," learning how to classify "pets" and "not pets" would have no effect on the decisions that the child makes about cats and dogs. Even though the child could interact with the same set of items, the attributes that are diagnostic of an item's "petness" would only affect the representation the child has for "pet" and "not pet." Another possibility is that interacting with an item that is both a cat (or dog) and a pet (or not a pet) would result in some modification of the original category representations acquired.

The simplest modification of existing knowledge would be that the attributes that are diagnostic in terms of the second classification set would have increased saliency in both category representations. Even though size is not diagnostic of "cat" (tigers and lions are quite large whereas domestic short-hairs are not), it is diagnostic of being a "pet," so size would have increased salience in the "cat" representation following the learning of the second classification set. The knowledge gained about the items from classifying them into the second classification set would essentially intrude into the representations of the primary classification set. Work by Ross (1997, 1999) shows that further interaction with the members of a learned classification set can result in this type of modification of the category representations. Although this work was not done with cross-classifiable items, it indicates that there might be some effect of the secondary learning on previous knowledge.

Current models of inductive category learning do not clearly address issues related to updating category representations by means of cross-classification. They rely on corrective feedback that would be specific to the category label of interest. So, if a tiger (a cat) is incorrectly classified as a pet, there is no readily available mechanism which might adjust the learner's representation of "cat." The topic of formally modeling the effects of cross-classification are addressed in the General Discussion.

Current Experiments

The purpose of the current studies is to assess whether category learning, specifically the increased salience of diagnostic attributes due to classification learning, will affect previous knowledge. In the experiments, the subjects learned about fictional creatures that could be classified by two different means. The attributes that were critical to learn the first classification set were non-informative for the secondary distinction (e.g., the legs were diagnostic in one classification set, but non-diagnostic in the other). The subjects initially learned one classification set and were asked to make category membership judgements about a series of the creatures. Then, the subjects were told that they would learn about different categories of creatures. The subjects saw the same creatures during this second classification learning, however, classification decisions were made in terms of the second classification set. After a transfer task concerning the second classification set, the subjects were asked to make the same judgements about the creatures they had after the initial learning. The analyses focus on whether there is a systematic effect of the secondary classification learning on the decisions the subjects make about the initial set.

It is hypothesized that subjects will alter their decisions regarding the items following the secondary learning primarily as a result of the increased salience of the attributes diagnostic for the second classification set. This would represent an *intrusion effect*, the knowledge critical for the second classification set would intrude upon the category knowledge of the primary classification set.

Experiment 1

A simple cross-classification structure was used to determine whether decisions based on an initially learned classification set were affected by subsequent learning of an orthogonal classification set. Individual creatures belonged to both a biological category and a category determined by eating habits. The attributes that were diagnostic of the biological classification set (the body shape, body covering, and legs) were more perceptually salient than the attributes that were diagnostic for the second classification set (the eyes, the mouth, and the hands). This difference is based on the assumption that most biological classifications do depend on salient features (such as body shape), while other distinctions may be more subtle (such as the mouth indicating a food preference). Also, since the attributes which are diagnostic for the second classification set are less salient, they will be less likely to inadvertently intrude upon decisions made about the primary classification set.

A typicality rating task was used to test for intrusion effects. Following learning, subjects rated the typicality of items that were perfectly consistent with the items they had seen during study and items that were inconsistent in terms of one of the two classification sets. The analyses focus on the subject responses to the items that were inconsistent in terms of the classification set learned second. For example, the eyes of the creatures were non-diagnostic during the initial classification learning, but during the typicality rating task, subjects saw creatures with eyes that had never occurred with the given hand shape and mouth. Since the attributes manipulated in these items were non-diagnostic in the first classification set, they should not have much influence on the initial typicality ratings. If there is no intrusion effect following the learning of the second classification set, the final typicality rating task should show the same pattern of results. If there is an intrusion effect, the typicality ratings of the items inconsistent with the second classification set should show a drop compared to the initial ratings.

Method

Subjects Sixteen undergraduates at the University of Illinois participated in the experiment for class credit.

Design Subjects first learned to classify the biological classification set (the AB set), and subsequently classify on the basis eating habits (the CD set). The stimuli within each study block were presented randomly. The order of the items during the typicality rating tasks was also randomized.

Materials The stimuli for this experiment were drawings of fictional creatures which were presented on computers. The creatures varied along six binary attributes; the body shape (squat or elongated), body covering (scales or fur), leg type (long and skinny or short and squat), eye type (cat-like or bug-like), mouth type (large and toothy or small and pursed), and hand type (rounded and stubby or long and claw-like). The first three attributes were correlated in how they were specifically instantiated, as were the last three attributes. The first three attributes were diagnostic of the AB classification set, and the other three attributes were diagnostic of the CD classification set. The classification sets were orthogonal so that an equal number of the study items were "Rosk Newbs" (or "Surk Tolls") that ate "Meabs" (or "Plats"). Abstractly, the attributes could be instantiated with a value of 1 or 0. A creature that had the values 111xxx, would be a "Rosk Newb," 000xxx would be a "Surk Toll," xxx111 would be a "Meab Eater," and xxx000 would be a "Plat Eater." By cross-classifying the items, a creature could have values of 111111, 111000, 000111, or 000000. The study stimuli were created by slightly modifying the features that instantiated the attributes of these four subset prototypes. This way, each of creatures within a subset were easily differentiated from one another, but they also clearly shared basic features within the classification sets. Four of each subset were made for study, and all 16 items were seen during each study block.

The stimuli for the typicality rating tasks were created in a similar manner. The stimuli were either novel consistent items, AB inconsistent items, CD inconsistent items, or old items. The old items were simply eight stimuli (two from each category subset) selected from the study items. The eight novel consistent items were created in the same way as the study items. These creatures were perfectly good members of one of the four creature subsets. The 24 inconsistent items were created by swapping one of the attribute values. The AB inconsistent items had one of the attribute values that was diagnostic for the taxonomic classification set in conflict (such as 110111) while the CD inconsistent items had a conflicting value present for one the attributes that was diagnostic for the other classification set (such as 111101). There were 40 items total in the typicality rating task, and the same items were used in all typicality rating tasks.

Procedure Subjects were told that they would be participating in a category learning experiment and that they should learn as much as they could about each of the creatures as they would be asked to make judgements about the creatures once they had finished the learning. No mention was made of the cross-classification of the items. All of the following instructions and tasks took place on the computer.

Subjects were told a biologist had discovered some new creatures on an island and their job was to learn how to identify these creatures. For each study trial, a creature appeared centered on the screen with the two possible category labels offset from center just below the picture. The subject clicked the mouse on either of the labels, and feedback on the classification was immediately given by the appearance of either "Correct" (in green) or "Incorrect" (in red) on both sides of the picture. The correct category label also appeared centered beneath the picture, and the subject could study the picture along with the category label as long as necessary. Each subject continued until a learning criterion of 15 out of 16 correct in a block was met and at least two blocks of study had been completed.

Once the learning criterion was met, instructions for the initial typicality rating task (AB1) were given. Items in the task were presented individually, blocked by the category of creatures. Beneath each picture was the question, "How typical is this creature of a Rosk Newb (or Surk Toll)?" along with a rating scale going from one ("Not at all Typical") to seven ("Very Typical"). The subject used the mouse to indicate the typicality rating, and then the next item appeared. After rating the typicality of items in both categories of the AB set, subjects were given instructions for the second classification set. Subjects were told that the biologist had discovered some new creatures, and the task was to learn how to identify the two types. The subjects were not told that they would be seeing the same creatures during this second classification task, although they most likely realized that fact once the second classification task began. The classification task and typicality rating task proceeded just as before, except that all decisions were made with regard to the CD classification set.

After completing the second typicality rating task, the subjects were informed about one last task – rating the typicality of a series of creatures in terms of the AB classification set. The final typicality rating task (AB2) was identical to AB1.

Results and Discussion

Did the typicality ratings of items change from AB1 to AB2? The analyses focus on the difference between ratings of the consistent items and the CD inconsistent items. The ratings of consistent items provide a baseline typicality for creatures that match the study items. By subtracting the mean typicality rating of the CD inconsistent items from that of the consistent items, the effect of the inconsistency among the attributes diagnostic for the second classification set is determined. If these features become more salient as a result of learning the CD classification set (even though they are not diagnostic in terms of the AB classification set), the drop in the typicality ratings for the CD inconsistent items should increase from AB1 to AB2. This change in the difference score is referred to as the *intrusion* *effect score*. All effects reported as significant are $p < .05$, two-tailed.

The mean typicality ratings for the items of interest are shown in Table 1. When the subjects had learned only about the AB classification set, there was a significant difference between the consistent items and the CD inconsistent items, $m = 0.34$, $SD = 0.46$, $t(15) = 2.92$. Although not intended, the attributes diagnostic for the CD classification set were apparently salient enough that subjects showed an effect of inconsistencies among those attributes. This effect was marginal in AB2, $m = 0.59$, $SD = 1.09$, $t(15) = 2.17$, because of the greater variability. The important question is whether there was a change in the level of the effect from AB1 to AB2. A significant increase would support the hypothesis that learning the CD classification set increased the salience of the attributes diagnostic in that set, and this affected the representations of the categories learned in the AB classification set. The intrusion effect score was positive, but it was not significant, $m = 0.26$, $SD = 1.09$, $t(15) < 1$.

Table 1: Mean Typicality Ratings by Item Type

	Experiment 1		Experiment 1b	
	AB1	AB2	AB1	AB2
Consistent	5.51	5.90	5.38	5.97
CD Inconsistent	5.17	5.31	4.92	5.34
Difference	0.34	0.59	0.45	0.63

Note: AB1 = initial typicality rating task, AB2 = final typicality rating task.

Experiment 1b

The intrusion effect score found in Experiment 1a was in the direction hypothesized, but non-significant. It is possible that the unlimited response time the subjects had in the typicality rating task was masking the effect. The attributes diagnostic for the CD classification set may have been more salient during AB2, as predicted by the intrusion effect, but the subjects had enough time to strategically ignore that source of information when making their typicality ratings. In this way, the intrusion effect score would have been diminished. To test for this, the rating task was modified so that the creature being rated was only briefly presented, reducing strategic effects.

Sixteen subjects were tested in a new version of Experiment 1a. All materials were the same. The subjects were told that the stimulus presentation in the rating task would be short, and during the task, the creatures were displayed for only one second following a 500ms fixation point.

Results and Discussion

The results of Experiment 1b were very similar to Experiment 1a. The modification of the typicality rating task did not increase the intrusion effect as expected. In AB1, the difference between ratings for the consistent items and the CD inconsistent items was significant, $m = 0.45$, $SD = 0.69$, $t(15) = 2.63$. The difference was also significant in AB2, $m = 0.63$, $SD = 1.02$, $t(15) = 2.48$, but the intrusion effect score ($m = 0.18$, $SD = 1.26$) was not significant, $t(15) < 1$.

As before, there was an increase in the difference scores from AB1 to AB2, but this increase did not indicate a reliable intrusion effect.

Experiment 1 Summary and Discussion

The initial analyses of data from Experiment 1a and 1b do not support the idea that the learning of an orthogonal classification set for a group of items has an effect on a previously learned classification set. However, in further analyses, a pattern was noticed between the AB1 difference scores and the intrusion effect scores. Small difference scores in AB1 occurred with positive intrusion effect scores, and larger difference scores in AB1 occurred with negative intrusion effect scores. A number of the subjects showed little sensitivity to the violations in the CD inconsistent items during AB1, which was what we had expected, and those subjects also showed the hypothesized intrusion effect. This was similar to the subjects in the studies by Ross (1997,1999). During the initial learning in those studies, the subjects were not sensitive to the use relevant attributes. The later use of the categories resulted in those attributes becoming more salient, much like what was expected by the intrusion effect.

So, for those subjects who were not sensitive to the non-diagnostic attributes following learning the AB classification set, there is evidence of the predicted intrusion effect. However, the subjects that were sensitive to the non-diagnostic attributes in AB1 show a qualitatively different effect. The typicality ratings for the CD inconsistent items increased after learning the CD classification set, the opposite of what was predicted by the intrusion effect. It appears that learning the CD classification set allowed these subjects to recognize what attributes were non-critical for the AB classification set, and they adjusted later decisions made about members of the initial categories accordingly. This strategic effect can be described as a *specification effect*. The subsequent learning helped to specify what was critical, and what was not critical, information in terms of the AB classification set.

All of this is highly speculative, but the occurrence of a specification effect would be important if true. If the previous summary is accurate, and the variation in the subject responses is not simply due to a lack of effect of learning the CD classification set, we should be able to make more specific predictions about the specification effect and when it is likely to occur. The second experiment is a test of the conclusions we drew from the post hoc exploration of the results of Experiment 1a and 1b.

Experiment 2

According to the proposed interpretation of the results of the first experiments, it should be possible to influence the effect score by manipulating the salience of the attributes diagnostic for the second classification set. To accomplish this, we had group of subjects learn the CD classification set first. Since the features for the AB classification set were more perceptually salient, they would more likely be incorporated into the initial learning of the CD classification set. The difference between the typicality ratings for the consistent items and the AB inconsistent items during the first typicality rating task should be large, even though the violations found within the AB inconsistent items occur within the non-diagnostic attributes. Once the subject has the opportunity to learn about the AB classification set, the secondary learning in this experiment, the subject would be more likely to modify his or her knowledge of the CD classification set by discounting the importance of attributes diagnostic of the AB classification set. The effect score should tend to be negative since the difference between the typicality ratings for the consistent items and the AB inconsistent items would decrease. Notice that this is contrary to the hypothesis initially proposed. If the negative effect score is found, it would provide support for the occurrence of a specification effect.

Method

The materials and procedures used in this experiment were identical to Experiment 1. The crucial difference was that the 16 subjects learned the CD classification set first and then learned the AB classification set.

Results and Discussion

The difference scores and intrusion effect score for each subject in Experiment 2 were determined as in Experiment 1. The mean typicality ratings for the items of interest can be found in Table 2. The first typicality rating task showed a large difference between the ratings for the consistent items and the AB inconsistent items, $m = 1.69$, $SD = 1.16$, and this difference was greater than zero, $t(15) = 5.84$. The difference in the final rating task was smaller, but still greater than zero, $m = 1.21$, $SD = 1.31$, $t(15) = 3.71$. The mean intrusion effect score for Experiment 2 was -0.48 ($SD = 0.89$), which was both different from the mean intrusion effect score found in Experiment 1 ($m = 0.26$), $t(30) = 3.12$, and different from zero, $t(15) = 2.16$. Whereas just over half of the subjects in Experiments 1a and 1b showed positive intrusion effect scores, only four of the sixteen subjects showed a positive intrusion effect score in Experiment 2. There is a clear tendency for the effect of violations to the diagnostic attributes for the second classification set to be reduced as result of learning the cross-classification. The predicted evidence for the specification effect was found.

When the CD classification set was learned first, the attributes for the AB classification set, although non-diagnostic, were salient. The AB inconsistent items, which contained violations to those attributes, were initially rated as being not very typical in terms of the CD classification set. After the subjects learned about the AB classification set, they adjusted how much they considered the attributes diagnostic of the AB classification set when rating the items in terms of CD classification set. The subjects were more able to specify what information was important when dealing with members of the CD classification set. The majority of the subjects in this experiment showed this specification effect because of how available the information about the AB diagnostic attributes was during the initial learning.

Table 2: Mean Typicality Ratings by Item Type from Experiment 2

	CD1	CD2
Consistent	5.83	5.95
AB Inconsistent	4.14	4.74
Difference	1.69	1.21

Note: CD1 = initial typicality rating task, CD2 = final typicality rating task.

General Discussion

The current study was intended to begin an investigation into an important issue, how category learning might affect prior knowledge. Although an important issue, it has received relatively little study, and past work indicates that once established, knowledge is fairly resistant to modification (Lewandowsky, Kalish, & Griffiths, 2000; Potts, St. John, & Kirson, 1989). Since the items we deal with on a daily basis can be considered to be members of more than one category, this cross-classification structure seemed to be an especially appropriate means to explore the possible effects. There has been little research into complex non-hierarchical category structures, and almost none dealing directly with category learning. This direction in category learning research is important since it reflects structure found in the real world, and because it raises important theoretical issues. The work that has been done showing the effect of prior knowledge on category learning has allowed for great gains to be made in our understanding of how intelligent systems learn. There is a great deal more to be learned as we come understand how current learning modifies previous knowledge.

The current research was intended to establish a methodology for investigating whether category learning leads to systematic changes in prior knowledge. Although the results are not entirely clear, the initial experiments were successful in challenging our predictions as to the possible effects of learning in a cross-classification structure. Based primarily on the results of Ross (1997, 1999), we expected to find evidence of an intrusion effect. Instead, we found evidence for two means by which category learning can affect prior knowledge. The results of Experiments 1 and 1b indicated that some subjects showed the predicted increase in salience for attributes diagnostic for the secondary classification set after learning to classify in terms of that classification set. However, other subjects showed a contrary tendency to reduce the importance of those attributes following the secondary learning. The second experiment showed further evidence of this unanticipated specification effect.

Both the intrusion and specification effects represent the interaction of knowledge from multiple categories for a given item. It is possible that the secondary learning directly causes the modification of preexisting category knowledge. However, it is also quite possible that the effects on the typicality ratings are the result of the synthesis of available information during the decision making process. This initial study is not able to differentiate between these possibilities. Still, there has been work both in the domain of machine learning (Dejong & Mooney, 1986) and cognitive psychology (Murphy & Allopenna, 1994; Wisniewski & Medin, 1994) that illustrates how prior knowledge constrains current learning because the information being learned is situated within the framework of knowledge already available. This process explains how prior knowledge affects learning, and it may also be a way to explain how prior knowledge is modified by the learning. If this is the case, the interaction between prior knowledge and learning is occurring sooner as opposed to later, but further study is necessary.

Most current models of category learning are not designed to account for the cross-classification of items (although see OLOC; Martin & Billman, 1994). Because of this, it is difficult to determine how applicable various models might be to the issue or how easily they might be modified in order to account for the effects found. The effects could be the result of a simple attribute weighting mechanism (like that found in ALCOVE; Kruschke, 1992). The intrusion effect could be the result of the increased saliency of attributes critical for the secondary classification set, and if the same mechanism was able to decrease the saliency of a given attribute as the result of learning a second classification set, it could also be considered an explanation for the specification effect. However, there would have to be some other process that informs the attribute weighting decision, and it is unclear what would fill this role. More recent models of categorization may have different solutions. SUSTAIN (Love, Markman, & Yamauchi, 2000) partitions categories in ways that would allow for subcategories to form, but again it is unclear how this might account for the contradictory effects. KRES (Rehder & Murphy, in press) offers another approach by placing knowledge modules into the system to guide the category learning. Although designed to account for knowledge effects on learning, there is a means to turn off unhelpful prior knowledge. With some modification, this could represent a mechanism for the specification effect. However these effects are modeled, it would be important for the mechanisms responsible for shifts in knowledge to be intimately connected to the knowledge itself, as proposed by Wisniewski and Medin (1994).

As people learn about categories, they face the challenge developing a sense of what is important among a grouping of items and how those items relate to other items and groupings of items. Category learning research has tended to not use learning situations where there are complex systems of relations present among items and categories. As a result, it has been difficult to gain an understanding of how it is that people develop dynamic and flexible knowledge structures.

Acknowledgments

Thanks to Kristine Onishi for helpful comments. This work was supported by NSF Grant SBR 97-20304.

References

Anderson, J. R. (1991). The adaptive nature of human categorization. *Psychological Review, 98*, 1409-1429.

Chin-Parker, S & Ross, B. H. (2002). Diagnosticity in category learning by classification and inference. In W. D. Gray & C. D. Schunn (Eds.), *Proceedings of the 24th Annual Conference of the Cognitive Science Society.* Mahwah, NJ: Erlbaum.

DeJong, G. F. & Mooney, R. (1986) Explanation-based learning: An alternative view. *Machine Learning, 1*, 145-176.

Heit, E. (2001). Background knowledge and models of categorization. In *Similarity and Categorization* (eds. Hahn and Ramscar), 155-178. Oxford University Press.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review, 99*, 1922-1944.

Lewandowsky, S., Kalish, M., & Griffiths, T. L. (2000). Competing strategies in categorization: Expediency and resistance to knowledge restructuring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1666-1684.

Love, B. C., Markman, A. B., & Yamauchi, T. (2000). *Modeling Classification and Inference Learning.* Paper presented at the Seventeenth National Conference on Artificial Intelligence (AAAI - 2000), Austin, TX.

Martin, J. D. & Billman, D. O. (1994). Acquiring and combining overlapping concepts. *Machine Learning, 16*, 121-155.

Murphy, G. L. (2002). *The Big Book of Concepts.* Cambrige, MA; MIT Press.

Murphy, G. L., & Allopenna, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 20*, 1904-1919.

Potts, G. R., St. John, M. F., & Kirson, D. (1989). Incorporating new information into existing world knowledge. *Cognitive Psychology, 21*, 1303-1333.

Rehder, B. & Murphy, G. L. (in press). A knowledge-resonance (KRES) model of knowledge-based category learning. *Psychonomic Bulletin and Review.*

Ross, B. H. (1997). The use of categories affects classification. *Journal of Memory and Language, 37*, 1240-1267.

Ross, B. H. (1999). Postclassification category use: The effects of learning to use categories after learning to classify. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1743-1757.

Ross, B. H., & Murphy, G. L. (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology, 38*, 1495-1553.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.

Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, *18*, 1221-1281.