Effects of Training on Category Learning

Wei (Sophia) Deng (deng.69@osu.edu)
Department of Psychology, The Ohio State University
267 Psychology Building, 1835 Neil Avenue
Columbus, OH 43210 USA

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
Department of Psychology, The Ohio State University
239 Psychology Building, 1835 Neil Avenue
Columbus, OH 43210 USA

Abstract

What information do people extract in the course of category learning? And how does training affect this process? The current study addressed these questions by examining the effects of training on the outcome of category learning in 4- to 5-year-olds and adults. In two experiments, participants were trained on either a classification task or an inference task and then tested with categorization and recognition tasks. The categorical information (i.e., deterministic and probabilistic features) was explicitly given to participants in Experiment 1 but not in Experiment 2. Results with adults replicate previous findings indicating that participants form different representations in the course of classification and inference training (rule-based representation in the former case and similarity-based representations in the latter case). In contrast, regardless of the type of training, young children form similarity-based representations.

Keywords: Cognitive Development, Categorization, Learning, Psychology.

Introduction

The ability to form categories is an important component of human cognition (see Murphy, 2002, for a review). It has been well established that this ability appears early in development, with young infants capable of forming categories (Eimas & Quinn, 1994; Oakes, Madole, & Cohen, 1991). The study of how categories are learned and used can elucidate "a single main theme to cognitive science – the question of how people come to have knowledge" (Murphy, 2002, p. 272).

The relationship between category learning and use can be examined by contrasting two of the fundamental functions of categories – classification and inference (E. Smith, 1994). To test theories of categorization, researchers developed a variety of tasks (see A. Markman & Ross, 2003, for a review), most of which are based on classification. In a typical classification learning task, participants are presented with stimuli, whose category membership is unknown, and are asked to predict a category each item belongs to. This situation is similar to that of sorting a set of squirrels and hamsters into two distinct groups. Whereas classification involves predicting the category of an item, inference involves predicting a missing feature using information from other features as well as the category. In this case, instead of determining whether an animal is a squirrel or a hamster, participants predict a value of a given feature (e.g., the type of tail the animal has).

There is evidence that classification and inference learners result in different representations of categories and much of these findings stem from a paradigm developed by Yamauchi and A. Markman (1998). The paradigm is based on the following idea. Imagine two categories A (labeled "A") and B (labeled "B"), each having four binary dimensions (e.g., Size: large vs. small, Color: black vs. white, Shape: square vs. circle, and Texture: smooth vs. rough). The prototype of Category A has all values denoted by "1" (i.e., "A", 1, 1, 1, 1) and the prototype of Category B has all values denoted by "0" (i.e., "B", 0, 0, 0, 0). There are two inter-related generalization tasks – classification and inference. The goal of classification is to infer category membership (and hence the label) on the basis of presented features. For example, participants are presented with all the values for an item (e.g., ?, 0, 1, 1, 1) and have to predict category label "A" or "B". In contrast, in the inference task participants have to infer a feature on the basis of category label and other presented features. For example, given an item (e.g., "A", 1, ?, 1, 0), participants have to predict the value of the missing feature. It was found that inference learners were more likely than classification learners to infer prototypical features which were correspondingly associated with training items. Multiple studies using this paradigm found that classification learners are sensitive primarily to diagnostic features that distinguish between categories, whereas inference learners are also sensitive to within-category correlations of features, which are not diagnostic but prototypical (Chin-Parker & Ross, 2002; Chin-Parker & Ross, 2004; Sakamoto & Love, 2006; Yamauchi, Love, & A. Markman, 2002; Yamauchi & A. Markman, 2000a, 2000b). Furthermore, there is also evidence that adults trained on a classification task attend to the most relevant features (A. L. Anderson et al., 2002). Adults learn to optimize performance in category learning by shifting their attention to different diagnostic features in different situations (Nosofsky, 1984; Rehder & Hoffman, 2005; Shepard et al., 1961) or learn inattention to newly relevant features (Hoffman & Rehder, 2010).

The argument that classification learning focuses on the diagnostic features distinguishing categories whereas inference learning focuses on the prototypical features reflecting within-category information is consistent with the evidence that adults' categorization is often rule-based (Rips, 1989; Allen & Brooks, 1991). Nosofsky and colleagues (Nosofsky & Palmeri, 1998; Nosofsky, Palmeri, McKinley, 1994) have proposed a quantitative model of human concept learning that learns to classify objects by forming simple logical rules and remembering occasional exceptions to those rules.

However, there is little agreement on the categorization process in early development. According to knowledge-based approaches, early in development, categorization and inductive generalization is considered to be based on prior categorical knowledge thus to be category based (Gelman & E. Markman, 1986; Gelman & Heyman, 1999; Gelman 2004). According to another approach, early categorization is similarity-based (Sloutsky & Fisher, 2004; Sloutsky, Kloos, & Fisher, 2007). There is evidence that early generalization is often driven by appearance similarity (Gelman, 1988; Gelman & E. Markman, 1987; Sloutsky & Fisher, 2004). In particular, infants are more likely to group items together if the items have overlapping within-group distributions of properties and non-overlapping between-group distributions (French, et al., 2004, see also Mareschal & Quinn, 2001; Mareschal, Quinn, & French, 2002). Similarly, infants are more likely to generalize non-obvious properties when the two items look alike (Welder & Graham, 2001) and they are more likely to extend a name to items that have similar shape (E. Markman & Hutchinson, 1984; L. Smith, et al., 1996). Similar results have been reported with young children, with similarity supporting both categorization of items and induction of non-obvious properties (Gelman, 1988; Gelman & E. Markman, 1987; Sloutsky & Fisher, 2004; Sloutsky & Lo, 1999).

Do children and adults show the same pattern of extracting and processing categorical information in category learning? How does training affect this process? And how do these effects change in the course of development? Does the asymmetry between classification and inference learning found in adults exist in children? Finding such an asymmetry would support the idea that, like adults, children treat classification and inference learning differently and tend to detect and rely on a defining feature to categorize items, whereas a symmetric performance in the classification and inference training would support the idea that children may perform similarity-based categorization and treat two types of category learning equally. The primary goal of this study is to address these questions.

Overview of the Current Study

Experiments reported here explored how categories were learned and used under classification and inference training by adults and young children. The basic task consisted of two phases, a training phase and a testing phase. During the training phase, participants had to infer either the category of a given item (in classification training) or a feature that the item has (in inference training). The testing phase consisted of categorization and recognition tasks and was administered immediately after the training phase. During the testing phase, which was identical for two training conditions, adult and child participants were asked to determine (1) which category the creature was more likely to belong to and (2) whether each picture was old or new. The structures of both training and testing stimuli will be described in the section below.

Experiment 1

Method

Participants There were 35 adults (16 women) and 21 preschool children (*M* = 56.6 months, range 53.2–59.5 months; 13 girls) participating in this experiment. In this and the second experiment reported here, adult participants were undergraduate students from the Ohio State University participating for course credit and were tested in a quiet room in our lab on campus. Child participants were recruited from childcare centers, located in middle-class suburbs of Columbus and were tested in a quiet room in their preschool by a female experimenter.

Materials In both experiments reported here, the materials, similar to those used previously by Deng and Sloutsky (2012, 2013), consisted of colorful drawings of artificial creatures that varied in their appearance and in a category-inclusion rule and that were accompanied by the novel labels "flurp" (Category F) and "jalet" (Category J). For these two categories, we created two prototypes (F0 and J0, respectively) that were distinct in the color and shape of seven of their features: head, body, hands, feet, antennae, tail, and button (see Figure 1). Two categories have a family-resemblance structure and stimuli were derived from the two prototypes by modifying the values of the seven features. The button is the deterministic feature (hereafter "D") and defines the category-inclusion rule: all members of Category F have raindrop-shaped button with the value of 1 whereas all members of Category J have cross-shaped button with the value of 0. All the other varying features – the head, body, hands, feet, antennae, and tail – constitute the probabilistic features (hereafter "P") and reflect the overall similarity among the exemplars.

The training stimuli consisted of High-Match items (i.e., $P_{flurp}D_{flurp}$ and $P_{jalet}D_{jalet}$). All members of $P_{flurp}D_{flurp}$ items had four probabilistic features (P) consistent with the prototype F0 with the value of 1 and two features consistent with the prototype J0 with the value of 0. And all of them have the deterministic feature (D) consistent with F0 valued 1. However, all members of $P_{jalet}D_{jalet}$ items had four probabilistic features consistent with the prototype J0 with the value of 0 and two features consistent with the prototype F0 with the value of 1. And all of them have the deterministic feature consistent with J0 valued 0.

The testing stimuli consisted of another four sets of items besides the High-Match items. The data analyses reported

Table 1. Example of category structure used in Experiment 1 and Experiment 2.

	Category F								Category J						
	Head	Body	Hands	Feet	Antenna	Tail	Button		Head	Body	Hands	Feet	Antenna	Tail	Button
F0	1	1	1	1	1	1	1	J0	0	0	0	0	0	0	0
$P_{flurp}D_{flurp}$	1	1	1	1	0	0	1	$P_{jalet}D_{jalet}$	0	0	0	0	1	1	0
$P_{jalet}D_{flurp}$	0	1	0	1	0	0	1	$P_{flurp}D_{jalet}$	1	0	1	0	1	1	0
$P_{all\text{-}new}D_{flurp}$	N	N	N	N	N	N	1	$P_{all\text{-}new}D_{jalet}$	N	N	N	N	N	N	0

Note. The value 1 = any of seven dimensions identical to Category F (flurp, see Figure 1). The value 0 = any of seven dimensions identical to Category J (jalet, see Figure 1). The value N = new feature which is not presented during training. P = probabilistic feature; D = deterministic feature. F0 is the prototype of Category F and J0 is the prototype of Category J.

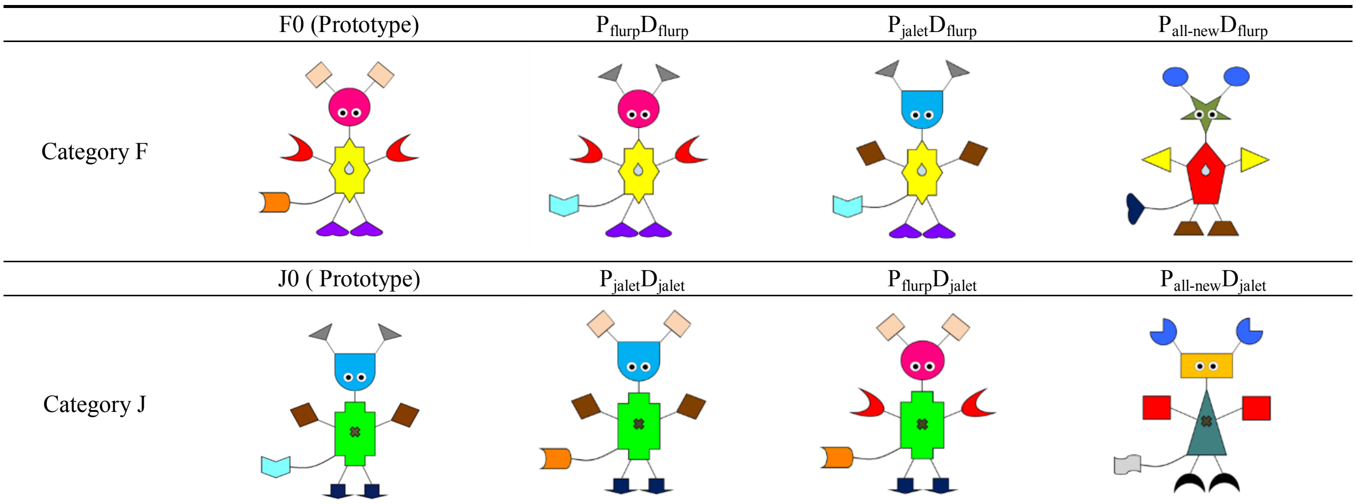

Figure 1. Examples of Stimuli Used in this study.

here only focused on two of them: critical lures (i.e., $P_{jalet}D_{flurp}$ and $P_{flurp}D_{jalet}$), and all-new-P items (i.e., $P_{all\text{-}new}D_{flurp}$ and $P_{all\text{-}new}D_{jalet}$). The High-Match items were used to examine participants' performance of category learning and to assess their recognition accuracy on the old items. Children were above 91.1% categorization accuracy on these trials, and adults were above 95.3%, all above chance (*ps* < .05) and exhibited memory accuracy of 82.2% and 94.5%, respectively. The all-new-P items were catch trials and consisted of six new probabilistic features which were not shown during training. Children and adults exhibited memory accuracy of 91.0% and 98.1%, respectively. The critical lures were Low-Match items: Most of the members of $P_{jalet}D_{flurp}$ items have the P with the value of 0 but all of them have D valued 1; whereas most of the members of $P_{flurp}D_{jalet}$ items have P with the value of 1 but all of them have D valued 0. This set of items was used to assess whether participants relied on overall similarity or category-inclusion rule to categorize new items. Table 1 shows example of category structure with P and D being combined to create three types of stimuli, and Figure 1 shows examples of each kind of stimulus.

Procedure The procedure consisted of two phases, a training phase and a testing phase. During the training phase, participants were given 30 trials (15 trials per category) and they had to infer either the category of a given item (in classification training) or a feature that the item has (in inference training). Each training trial was accompanied by corrective feedback. The classification and inference training differed in the type of dimensions being predicted. In classification training, participants predicted the category label of a stimulus given information about all other features. In inference training, participants predicted one missing feature of a stimulus given the information about the remaining features as well as the label. The information about P and D was explicitly given to participants before training. They were told that all flurps (or jalets) had raindrop button (or cross button) and most of them had flurps' (or jalets') P by presenting corresponding probabilistic features one at a time. This information was repeated in the corrective feedback to each response during training. Adult and child participants were randomly assigned to one of the two training conditions. The testing tasks were not mentioned in the training phase of any of the conditions.

Testing phase, including categorization and recognition tasks, was administered immediately after training. During the testing phase, which was identical for two training conditions, adult and child participants were presented with 40 trials of creatures and were asked to determine (1) which category the creature was more likely to belong to and (2) whether each picture was old (i.e., exactly the one presented during the training phase) or new. The order of the 40 items was randomized across participants. No feedback was provided during the testing phase.

The procedures were identical for both adult and child participants except the way the instructions were presented and the questions were asked. Adult participants read the instructions and questions on the computer screen and pressed the keyboard to make responses, whereas for children instructions as well as questions were presented verbally by a trained experimenter and the experimenter recorded children's responses by pressing the keyboard. The proportion of responses in accordance with the category from which the exemplar was derived (i.e., rule-based responses) was the dependent variable. If classification learners and inference learners process and represent categorical information differently, their performance should be asymmetric between Classification Training and Inference Training conditions. However, if there is no difference between classification and inference training, participants should show symmetric pattern between two training conditions. In addition, if participants rely on the deterministic feature, the proportion of rule-based responses should be high. However, if they rely on multiple probabilistic features, they should make low level of rule-based response.

Results and Discussion

All results reported here only focused on performance of the categorization task, specifically on the critical lures (i.e., $P_{flurp}D_{jalet}$ and $P_{jalet}D_{flurp}$ items). Recall that if participants form a rule-based representation of a category, they should identify the $P_{flurp}D_{jalet}$ item as a jalet, whereas if they formed a similarity-based representation, they should identify this item as a flurp.

The main results of Experiment 1 are shown in Figure 2. As shown in the figure, children tended to form similarity-based representations regardless of condition, whereas adults tended to form rule-based representations in the classification condition. These findings were supported by statistical analyses – data in the figure were analyzed with 2 (Training Type: Classification and Inference) by 2 (Age Group: 4-5-year-olds vs. Adults) between-subjects ANOVA. There was a main effect of training type, $F(1,52) = 5.56$, $MSE = 0.38$, $p = .022$, $\eta^2 = 0.097$, and a main effect of age group, $F(1,52) = 24.49$, $MSE = 1.65$, $p < .001$, $\eta^2 = 0.320$. Specifically, adults made more rule-based responses in Classification Training than in Inference Training, independent samples $t(1,31.2) = 2.63$, $p = .013$, $d = 0.92$, with the proportion of rule-based responses above chance in Classification Training, one-sample $t(1,15) = 4.28$, $p = .001$, $d = 1.07$, but around chance in Inference Training, one-sample $t(1,18) = 0.94$, $p = .359$. However, for children, they made comparable rule-based responses in both training conditions ($p = .334$), with the proportion of rule-based responses significantly below chance in Inference Training, one-sample $t(1,6) = 5.46$, $p = .002$, $d = 2.06$, and marginally below chance in Classification Training, one-sample $t(1,13) = 2.03$, $p =.064$, $d = 0.54$.

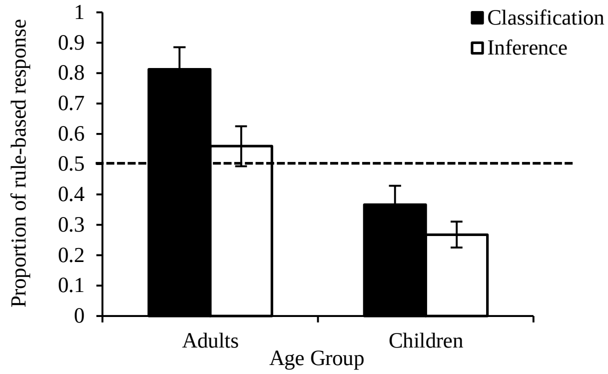

Figure 2. Proportion of rule-based responses by age group and training type in Experiment 1.

The results are consistent with previous evidence (Yamauchi & A.Markman, 1998; Hoffman & Rehder, 2010) pointing to the predicted asymmetry between classification and inference training for adults. As predicted, adults tended to process and represent categorical information differently, with classification learners being more likely than inference learners to focus on deterministic feature, which separates two categories. However, there was little evidence that children learned categories differently by classification and inference. The symmetric performance suggested that children treated classification training and inference training equally and, more importantly, unlike adults, children formed similarity-based representation of categories.

One possible limitation of Experiment 1 was that participants were told explicitly about the deterministic and probabilistic features. It is possible that only adults, but not children attended to this information, and as a result, only adults formed rule-based representations. Experiment 2 attempted to eliminate this possibility by not mentioning that there were probabilistic and deterministic features.

Experiment 2

Method

Participants Twenty-six adults (18 women) and twenty preschool children (M = 55.3 months, range 49.8–60.2 months; 7 girls) participated in this experiment. Two additional adults were texting during experiments and these data were excluded from the analysis.

Materials and procedure The materials were identical to those used in Experiment 1. The overall procedure in Experiment 2 was identical to Experiment 1 except that neither the information of P nor D was given to participants before main experiment or in the feedback (i.e., participants were only given corrective feedback). For the old items (i.e., High-Match items) at test, children were above 70.1% categorization accuracy on these trials, and adults were above 83.9%, all above chance (ps < .05) and exhibited memory accuracy of 75.0% and 71.1%, respectively. For the all-new-P items, Children and adults exhibited memory accuracy of 78.2% and 82.2%, respectively.

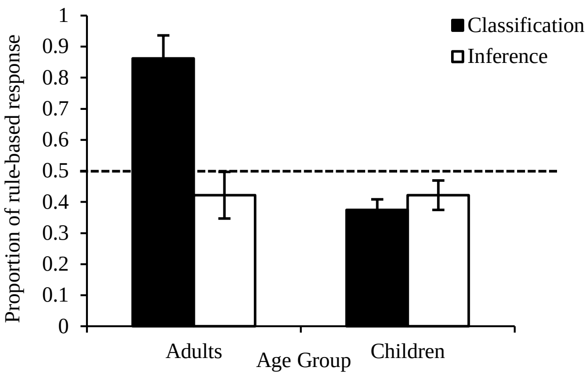

Figure 3. Proportion of rule-based responses by age group and training type in Experiment 2.

Results and Discussion

The main results of Experiment 2 are shown in Figure 3. The data were analyzed with 2 (Training Type: Classification and Inference) by 2 (Age Group: 4-5-year-olds vs. Adults) between-subjects ANOVA. The results revealed a significant training type by age group interaction, $F(1,42) = 16.48$, $MSE = 0.64$, $p = .001$, $\eta^2 = 0.282$. Independent samples t test indicated that adults made more rule-based responses in Classification Training than in Inference Training, $t(1,24) = 6.14$, $p = .001$, $d = 2.58$, with the proportion of rule-based responses above chance in Classification Training, one-sample $t(1,9) = 4.95$, $p = .001$, $d = 1.56$, but below chance in Inference Training, one-sample $t(1,15) = 2.30$, $p = .036$, $d = 0.57$. However, children exhibited comparable proportions of rule-based responses in both training conditions ($p = .604$), with the proportion of rule-based responses around chance in Classification Training ($p = .125$) and Inference Training ($p = .140$).

The results in Experiment 2 revealed the same pattern as Experiment 1. For adults, there was an asymmetry between classification and inference training; whereas young children's performance in the two training conditions was symmetric, and, regardless of the training condition, they formed similarity-based representations.

General Discussion

The reported study examines the effects of training on the outcome of category learning and changes in these effects in the course of development. To achieve this goal, we trained adult and child participants with a category learning task in which participants learned two categories. Each category had a single deterministic feature that differed between the categories and multiple probabilistic features that partially overlapped between categories. Participants who were trained on a classification task were asked to classify items into one of two categories; whereas participants who were trained on an inference task were asked to infer a missing feature of items. Following training, participants were tested on their ability to categorize novel items.

Two major findings stem from the reported results. First, in both reported experiments adults exhibited an asymmetric pattern between classification training and inference training. Their rule-based responses in classification training were consistently higher than those in inference training, which is consistent with previous evidence (Yamauchi & A. Markman, 1998; Hoffman & Rehder, 2010) suggesting that adults process and represent categorical information differently. Specifically, classification learners are more likely to focus on deterministic (or rule-based) features than inference learners. However, for young children, the symmetry between classification and inference suggests that they do not treat these two training conditions differently. Second, adults tend to spontaneously detect a defining feature (Experiment 2) and classification learners tend to consistently rely on it to categorize items (Experiment 1 and 2). But there is little evidence that children rely on the deterministic feature in categorization. In contrast to adults, children tend to rely on a pattern of correlated probabilistic features, which reflects the overall similarity.

The results have implications for understanding the mechanisms underlying category learning and how these mechanisms may change in the course of development. Future research will also examine how attention is allocated in category learning by using a combination of eye tracking and behavioral paradigm.

Acknowledgements

This research was supported by a National Science Foundation (NSF) grant (BCS-0720135) to Vladimir M. Sloutsky.

References

Allen, S. W., & Brooks, L. R. (1991). Specializing the operation of an explicit rule. *Journal of Experimental Psychology: General, 120,* 3-19.

Anderson, A. L., Ross, B. H., & Chin-Parker, S. (2002). A further investigation of category learning by inference. *Memory and Cognition, 30*, 119-128.

Chin-Parker, S., & Ross, B. H. (2002). The effect of category learning on sensitivity to within-category correlations. *Memory and Cognition, 3*, 353-362.

Chin-Parker, S., & Ross, B. H. (2004). Diagnosticity and prototypicality in category learning: A comparison of inference learning and classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 216-226.

Deng, W., & Sloutsky, V. M. (2012). Carrot-eaters and moving heads: Inductive Inference is better supported by salient features than by category labels. *Psychological Science, 23,* 178-186.

Deng, W., & Sloutsky, V. M. (2013). The role of linguistic labels in inductive generalization. *Journal of Experimental Child Psychology, 114,* 432-455.

Eimas, P. D., & Quinn, C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, *65,* 903-917.

French, R.M., Mareschal, D., Mermillod, M., & Quinn, P.C. (2004). The role of bottom-up processing in perceptual categorization by 3- to 4-month-old infants: Simulations and data. *Journal of Experimental Psychology: General*, *133,* 382-397.

Gelman, S.A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology*, *20*, 65-95.

Gelman, S.A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8,* 404-409.

Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science, 10,* 489-493.

Gelman, S. A., & Markman, E. (1986). Categories and induction in young children. *Cognition, 23*, 183-09.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58, 1532*-1541.

Hoffman, A. B., & Rehder, B. (2010). The costs of supervised classification: The effect of learning task on conceptual flexibility. *Journal of Experimental Psychology: General, 139,* 319-340.

Mareschal, D., & Quinn, P. C. (2001). Categorization in infancy. *Trends in Cognitive Sciences, 5,* 443-450.

Mareschal, D., Quinn, P.C., & French, R.M. (2002). Asymmetric interference in 3- to 4-month olds' sequential category learning. *Cognitive Science, 26,* 377-389.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16,* 1-27.

Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin, 129,* 592-613.

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

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 104-114.

Nosofsky, R. M., & Palmeri, T. J. (1998). A rule-plus-exception model for classifying objects in continuous-dimension spaces. *Psychonomic Bulletin & Review, 5*, 345-369.

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

Oakes, L. M., Madole, K. L., & Cohen, L. B. (1991). Object examining: Habituation and categorization. *Cognitive Development, 6,* 377-392.

Rehder, B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology*, *51*, 1-41.

Rips, L. J. (1989). Similarity, typicality, and categorization. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 21-59). New York: Cambridge University Press.

Sakamoto, Y., & Love, B. C. (2006). Vancouver, Toronto, Montreal, Austin: Enhanced oddball memory through differentiation, not isolation. *Psychonomic Bulletin & Review, 13,* 474-479.

Shapard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classification. *Psychological Monographs*, *75*(Whole No.517).

Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General, 133,* 166-188.

Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science*, *18,* 179-185.

Sloutsky, V. M., & Lo, Y.-F. (1999). How much does a shared name make things similar? Part 1: Linguistic labels and the development of similarity judgment. *Developmental Psychology, 35,* 1478-1492.

Smith, E. E. (1994). Concepts and categorization. In E. E. Smith & D. N. Osherson (Eds.), *An invitation to cognitive science* (pp. 3-33). Cambridge, MA: MIT Press.

Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition*, *60,* 143-171.

Welder, A. N., & Graham, S. A. (2001). The influences of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development, 72,* 1653-1673.

Yamauchi, T., Love, B. C., & Markman, A. B. (2002). Learning nonlinearly separable categories by inference and classification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 3,* 585-593.

Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory and Language*, *39,* 124-148.

Yamauchi, T., & Markman, A. B. (2000a). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 776-795.

Yamauchi, T., & Markman, A. B. (2000b). Learning categories composed of varying instances: The effect of classification, inference, and structural alignment. *Memory and Cognition, 28,* 64-78.