Effects of linguistic and perceptual information on categorization in young children

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
Center for Cognitive Science & School of Teaching & Learning
Ohio State University
21 Page Hall, 1810 College Road
Columbus, OH 43210, USA

Anna V. Fisher (fisher.449@osu.edu)
Center for Cognitive Science & School of Teaching & Learning
Ohio State University
21 Page Hall, 1810 College Road
Columbus, OH 43210, USA

Abstract

This paper examines the process of categorization in young children, and tests predictions derived from a model of young children's similarity judgment. The model suggests that linguistic labels might have greater contribution to similarity judgment for younger children than do other attributes. It is argued that because categorization is based on similarity, the model predicting similarity judgment should also predict categorization. Predictions of the model were tested in the experiment where 4-6 year-olds were asked to perform a categorization task. Results of the experiment demonstrate that young children perform categorization in a similarity-based manner, and support both qualitative and quantitative predictions of the label-as-attribute model.

Introduction

The ability to group things together is an important component of human cognition: stimuli (i.e., objects, scenes, situations, or problems) rarely recur exactly, and, as a result, records of specific stimuli would be of little help. Therefore, the ability to form categories and store stimuli as members of these categories is a critical component of learning, memory, and thinking. Furthermore, it has been demonstrated that even infants are capable of forming categories (e.g., Balaban & Waxman, 1997; Quinn & Eimas, 1998; Mandler, 1997). It is less clear, however, how people form categories and include novel instances into a category. Several theories have emerged in an attempt to answer these questions (e.g., Lamberts & Shanks, 1997; Smith & Medin, 1984, for reviews).

The general question of how people form categories and add new instances to these categories consists of three more specific questions: (1) How do people decide whether or not a novel entity is a member of an existing category? (2) How do people form a category when presented with a large number of positive and negative instances of the category? And (3) how do people decide whether or not two novel entities are members of the same novel category? While much theoretical and empirical work on categorization has focused on the first two questions (see Lamberts & Shanks, 1997; Smith & Medin, 1984, for reviews), the third question has remained largely under-researched. At the same time, answers to this question are important for understanding of the "first step" in the process of categorization – forming a new category and including some novel entities as its members, while excluding others.

The current research attempts to examine the third question. As a first approximation, it appears plausible that, if no information about the entities is available, the entities would be grouped together on the basis of their perceptual similarity. If, in addition to perceptual information, there is also linguistic information (e.g., "*Look, here is an X*"), then there are at least two possibilities for grouping. If the label *X* is familiar, then the object denoted as *X* could be included into all categories that include *X* as its member. However, if label *X* is novel, it seems likely that categorization should be performed on the basis of similarity. In this case, a model predicting similarity judgment should also predict categorization. One such model, the label-as-attribute model suggests that young children consider linguistic labels as attributes of compared entities (Sloutsky & Lo, 1999). The model predicts that both perceptual and linguistic cues should contribute to comparison-based processes, such as similarity judgment. These predictions have been confirmed in a number of studies examining contribution of perceptual and linguistic factors to similarity judgment (Sloutsky & Lo, 1999) and inductive inference (Sloutsky & Lo, 2000; Sloutsky, Lo, & Fisher, in press). It was found that young children aggregate perceptual and linguistic cues when computing overall similarity among compared entities.

We can predict, therefore, categorization should be a function of similarity computed over perceptual and linguistic cues. In what follows, we specify the model and its predictions, and present experiments designed to test predictions of the model.

The model is based on the product-rule model of similarity (Estes, 1994; Medin, 1975) that specifies similarity among non-labeled feature patterns. In the product-rule model, similarity is computed using Equation 1:

$$Sim(i, j) = S^{N-k} \qquad \mathbf{1}$$

where N denotes the total number of relevant attributes, k denotes the number of matches, and S ($0 \leq S \leq 1$) denotes values (weights) of a mismatch. For example, suppose that one is presented with two schematic faces A and B. Further suppose that these faces consist of four distinct features (i.e., the shape of the face, eyes, nose, and the size of ears), that they share two of these features (i.e., the shape of the face and eyes), and differ on the other two. Finally, suppose that $S = 0.5$, the value frequently derived empirically in past research (Estes, 1994). In this case, similarity between A and B would be equal to 0.25 (i.e., 0.5^2). Note that similarity between entities decreases very rapidly with a decrease in the number of mismatches, approximating the exponential decay function discussed elsewhere (Nosofsky, 1984). For example, if the faces share only one of the four features, their similarity would be equal to 0.125 (i.e., 0.5^3). On the other hand, if the faces share all four features, they would be identical, and their similarity would be equal to 1 (i.e., 0.5^0).

The label-as-attribute model suggest that linguistic labels might have greater contribution to similarity judgment for younger children than do other attributes, and there is evidence supporting this suggestion (see Sloutsky & Lo, 1999). Why would labels weigh more for younger children and what might be a mechanism underlying the greater weight of labels at earlier age? One possible explanation is that labels have larger weights because they are presented auditorily, and the auditory system matures earlier than the visual system. In particular, the auditory system starts functioning during the last trimester of gestation (Birnholz & Benaceraff, 1983; see also Jusczyk, 1998, for a review), whereas the visual system does not start functioning until after the birth. As a result, even though the neural bases of visual perception are fully developed at quite a young age (e.g., Aslin & Smith, 1988), auditory stimuli may still have a privileged processing status for younger children, thus resulting in larger weights of auditory stimuli (Napolitano, Sloutsky, & Boysen, 2001). In fact, it has been demonstrated that 15-months-olds grouped objects together when the objects shared an auditory input (either a label or a non-linguistic instrumental music input) if the input perfectly correlated with an infant's fixation of an object (Roberts, 1995; Roberts & Jacob, 1991, but see Balaban & Waxman, 1997).

According to the label-as-attribute model, similarity of labeled feature patterns could be calculated using Equation 2:

$$Sim(i, j) = S_{Label}^{1-L} S_{Vis.attr}^{N-k} \qquad \mathbf{2}$$

where N denotes the total number of visual attributes, k denotes the number of matches, $S_{vis.attr.}$ denotes values (attentional weights) of a mismatch on a visual attribute, S_{Label} denotes values of label mismatches, and L denotes a label match. When there is a label match, $L = 1$, and $S_{Label} = 1$; when there is a label mismatch, $L = 0$, and $S_{Label} < 1$. Note that S ($0 \leq S \leq 1$) denotes attentional weights of mismatches and the contribution of S is large if S is close to 0 and is small if S is close to 1. This is because the closer the value of S to 1, the smaller the contribution of a mismatch to the detection of difference, while the closer the value of S to 0, the greater its contribution to the detection of difference. When two entities are identical on all dimensions (i.e., there are no mismatches), their similarity should be equal to 1; otherwise, it is smaller than 1. Note that according to the model, when neither entity is labeled (i.e., $S_{Label} = 1$), similarity between entities is determined by the number of overlapping visual attributes, thus conforming to Equation 1. Labels are presented as a separate term in the equation because they are expected to have larger attentional weights than most visual attributes, an assumption that was borne out in previous research (Sloutsky & Lo, 1999). In the case that the weight of a label does not differ from that of other attributes, the label will become one of the attributes in the computation of similarity, and Equation 2 turns into Equation 1.

Finally, the model suggests that if the child is presented with a Target feature pattern (T) and Test feature patterns (A and B) and asked which of the Test patterns is more similar to the Target, the child's choices could be predicted using Equation 3:

$$P(B) = \frac{Sim(T,B)}{Sim(T,B) + Sim(T,A)} \qquad \mathbf{3}$$

In short, we argue that if categorization in young children is indeed similarity-based, then the same model that predicts similarity judgment in young children (e.g., Sloutsky & Lo, 1999) should be able to predict their categorization.

Simple derivations from Equation 3 allow us to predict categorization as a function of feature overlap. First, consider the case when entities are not labeled. Substituting *Sim (T,A)* and *Sim (T,B)* by their equivalents in Equation 1, we get Equation 4:

$$P(B) = S^x/(S^x + S^y) = S^x/[S^x(1 + S^{y-x})] = 1/(1 + S^y/ S^x) \quad \mathbf{4}$$

For the labeled entities, derivations remain essentially the same, except for the S_{Label} parameter. The parameter equals to 1, if there is a label match, otherwise it equals to λ ($0 < \lambda < 1$). Therefore, in the case of labeled entities, the probability of selecting the item that shared the same label (say item B) could be derived according to Equation 5:

$$P(B) = S^x/(S^x + \lambda S^y) = S^x/[S^x(1 + \lambda S^{y-x})] = 1/(1 + \lambda S^y/S^x) \quad \mathbf{5}$$

In short, in the no-label condition, the probability of categorizing Test B and the Target together should be a function of the ratio of S^y/S^x (i.e., the ratio of similarity of Test A and Test B to the Target), whereas in the label condition such categorization should be a joint function of S^y/S^x and λ (i.e., attentional weight of label). Because we can estimate λ from our prior research, we can use Equations 4 and 5 for estimating specific probabilities of categorization. One important (and testable) consequence of this proposal is that because linguistic labels contribute to similarity in a quantitative manner rather than in a qualitative "all-or-nothing" manner, they should also make a quantitative contribution to categorization.

When stimuli consist of a small number of easily distinguishable and countable features (e.g., schematic faces or dot patterns), *N* and *K* (Equations 1 and 2) and subsequently *X* and *Y* (Equations 4 and 5) could be computed directly. However, if stimuli are perceptually rich, the task of determining *N*, *K*, and subsequently *X* and *Y* is complicated, if not impossible. One possible solution to this problem is to conduct a calibration study estimating similarity of each of the Test stimuli to the Target. Because similarity of each of the two Test stimuli to the Target is equivalent to S^y and S^x, ratios of similarity (i.e., S^y/S^x) could be easily computed, and therefore could be used to test the model.

The overall experimental idea was as follows. Participants were presented with triads of stimuli, with each triad consisting of Test stimuli A and B and Target T. In order to use perceptually rich stimuli and to quantify perceptual similarity, the stimuli were selected from sequences of images, in which one animal was "morphed" into another in a fixed number of steps. An example of a morphed sequence is presented in Figure 1. Multiple triads were formed from these sequences. These triads were subjected to a preliminary "calibration" study, in which participants were asked to estimate similarity of each of the Test stimuli to the Target. Those triads that gave the ratios of .5/.5, .4/.6, .3/.7, and .1/.9 were selected for the major study. An example of a .3/.7 triad is presented in Figure 2.

In addition to the quantitative predictions of Equations 4 and 5, we can formulate two qualitative predictions:

(1) When entities are not labeled, the probability of categorizing of Test stimulus together with the Target is a function of the ratio of perceptual similarity of this Test stimulus and the competing test stimulus to the Target.

(2) When entities are labeled, linguistic labels should affect categorization in a quantitative manner rather than in a qualitative "all-or-nothing" manner. Categorization should be a function of two variables – the weight of linguistic label and the similarity ratios – and not of linguistic labels alone.

Method

Participants

Participants were 37 preschool children recruited from daycare centers located in middle class suburbs of Columbus, Ohio (19 girls and 18 boys, *M* = 5.4 years; *SD* = 0.82 years).

Materials and Design

The experiment had a mixed design with a labeling condition (label vs. no-label) as a between-subject variable and similarity ratio as a within-subject variable. At both levels of the labeling condition participants were presented with the same triads of animal faces, one of which was a Target and two of which were Test stimuli. The only difference between the levels of the labeling condition was that in the label condition all stimuli were labeled, whereas in the no-label condition these stimuli were not labeled.

Materials consisted of triads of 4" × 4" pictures of animal faces selected to represent four levels of the stimulus pattern condition. Selection was made on the basis of results obtained in the calibration study. Each triad of pictures included a Target and two Test stimuli. The Target was located at the center above the Test stimuli.

These stimuli were selected by conducting a calibration experiment, in which 19 4–5 year-old children were presented with triads of pictures of animals (similar to those in Figure2) and asked which of the Test stimuli was more similar to the Target. 16 triads were selected on the basis of this calibration, representing four similarity ratios of (e.g., *Sim* (A, Target)/ *Sim* (B, Target): (1) .5/.5 = 1, (2) .4/.6 = 1.5, (3) .3/.7 = 2.33, and (4) .1/.9 = 9. Each of the four ratios included 4 triads. These four levels of perceptual similarity were included in the design.

Figure 1. Examples of 5 steps in a 20-step morphing sequence.

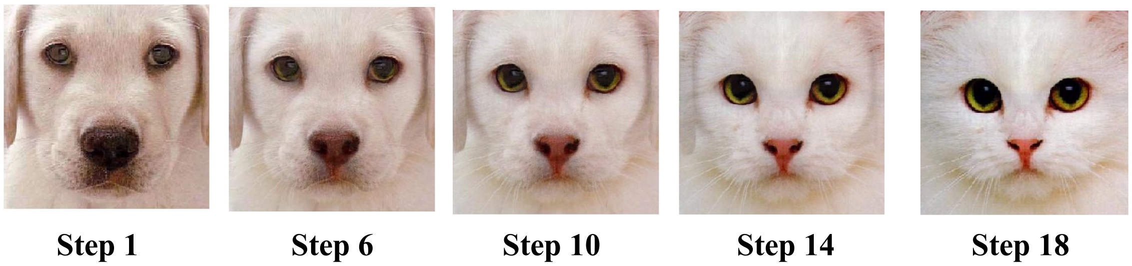
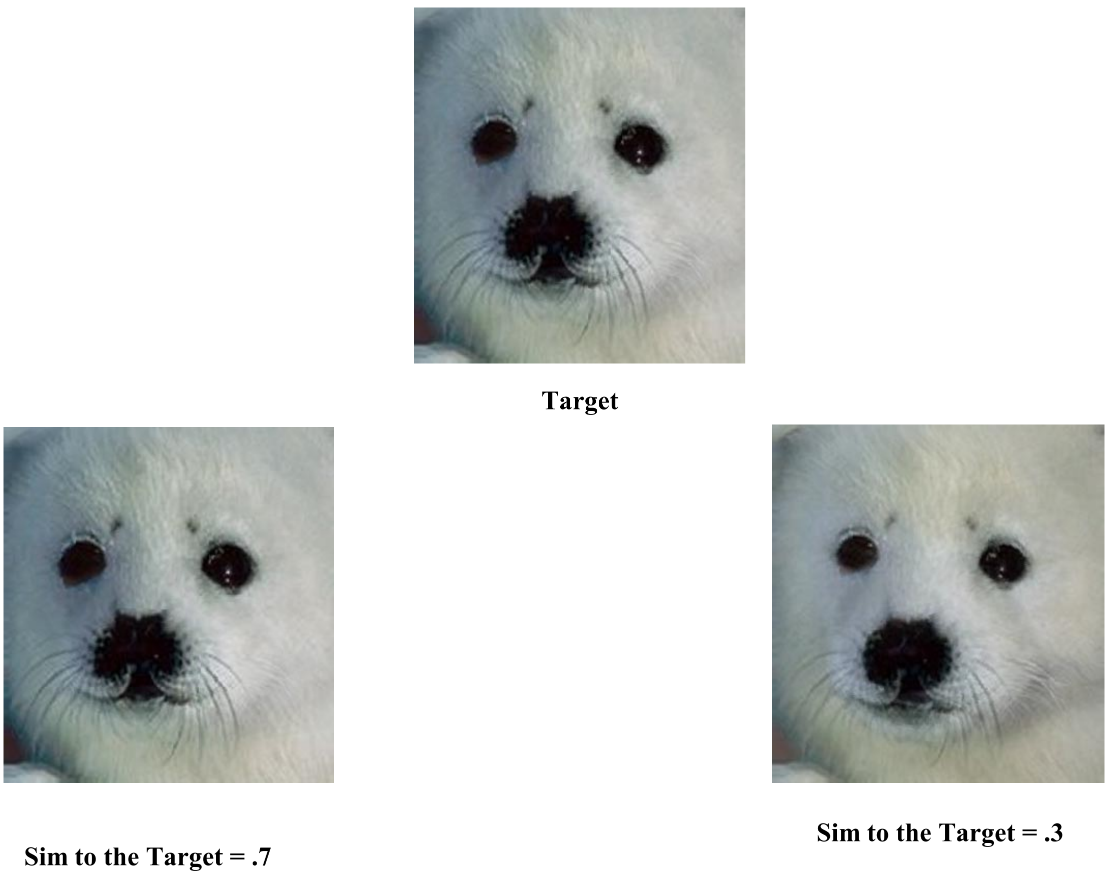


Figure 2. Example of an experimental triad



Procedure

Triads of pictures were presented to each participant on a computer screen. A female researcher interviewed each child individually in a quiet room in their schools. Before the experimental task participants were introduced to two warm-up trials. Questions asked during the warm-up trials were identical to the questions asked during the experimental trials. No feedback was given to the participants on their performance on the warm-up or experimental trials, and no participant was eliminated from the study on basis of his/her performance in the warm-up. The sole purpose of the warm-up was to illustrate to children the nature of the task they were to perform.

Experimental trials were identical to warm-up trials. In the label condition participants were first introduced to the labels for the Target and Test stimuli and asked to repeat them. All the labels used were two-syllable artificial count nouns (e.g. a Bala, a Guga). No labels were introduced in the no-label condition. Then, children were asked whether the Target was the same kind of animal as Test 1 or Test 2. Positions of two Test stimuli were randomized across trials. In both conditions participants had 16 experimental trials (four trials each of the four within-subject stimulus patterns). The order of trials was randomized for each participant.

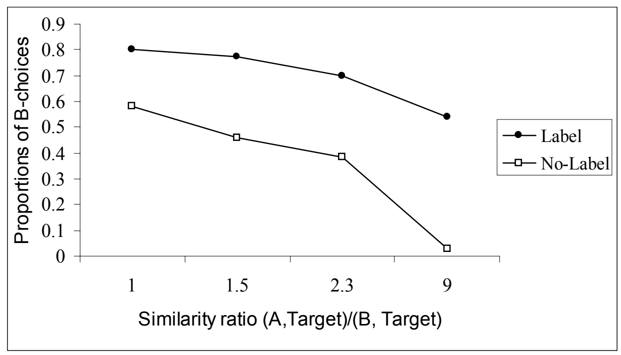
The important part of the instruction for preschool participants read: *Now we are going to play a game about animals from other planets. I am going to show you pictures of those alien animals, tell you their names, ask you to remember their names, and repeat them to me. Then I will ask you one question about those animals. Are you ready to start? I will show you something like this* (a warm-up triad was introduced at this point). *Look at them: this is a Guga* (points to the Target). *This is a Bala* (points to Test A). *This is a Guga* (points to Test B). *Could you please repeat their names? Do you think that this Guga* (points to the Target) *is the same kind as this Bala* (points to Test A) *or this Guga* (points to Test B)?

Note that in the no-label condition all stimuli were referred to as "this one." The order of introduction of the Test stimuli and their location relative to the target were randomized.

Results and Discussion

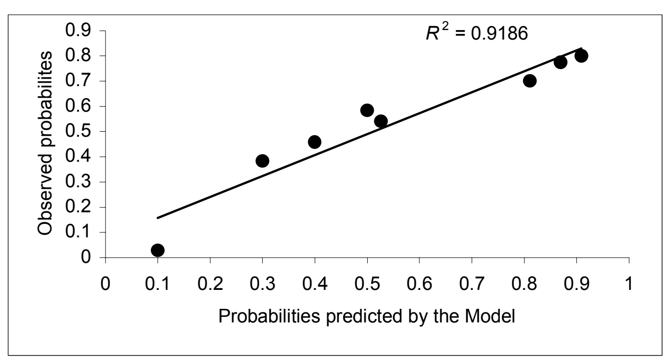
Proportions of B-choices by levels of the similarity ratio and labeling condition are presented in Figure 3 (recall that in the Label condition, Test B always shared the label with the Target). These proportions were averaged across the four trials for each level of the ratio and then averaged across subjects. Proportions averaged across trials were then subjected to a two-way (Labeling condition by Similarity ratio) mixed ANOVA with levels of similarity ratio as a repeated measure. The analyses indicated a significant main effect of labeling ($M_{\text{Label}} = .70 > M_{\text{No-Label}} = .36$), $F(1,35) = 30$, $MSE = 0.15$, $p < .0001$, and a significant main effect of the similarity ratio, $F(3,105) = 22.9$, $MSE = 0.06$, $p < .0001$, with no significant interaction. Planned comparisons of the levels of similarity ratio pointed to the following direction P(1) > P(1.5) = P(2.33) > P(9), all ts > 2, ps < .05. These results support the qualitative predictions, indicating that (a) when entities are not labeled, categorization is a function of perceptual similarity; (b) when entities are labeled, categorization is a function of similarity computed over perceptual and linguistic cues, and (c) labels contribute quantitatively to similarity among entities.

Figure 3. Proportions of B-choices by similarity ratio and labeling condition.



Quantitative predictions of the model are presented in Figure 4, where predicted probabilities of B-choices are plotted against observed probabilities. For the no-label condition, these probabilities were derived from Equation 4, whereas for the label condition they were derived from Equation 5 ($\lambda = .1$ was estimated from previous Sloutsky & Lo's data sets). Results indicate a good fit between predicted and observed probabilities ($r = .95$) with approximately 92% of variance explained by the model. These results indicating that similarity predicts much of categorization in young children support the hypothesis that, at least when labels are novel, categorization in young children is a function of similarity.

Figure 4. Overall fit of the model.



Several issues, however, would require further research. In particular, it remains unclear whether or not adults exhibit the same pattern of categorization as children. On the one hand, if entities are novel, it seems likely that adults should also use similarity as a basis of their categorization. On the other hand, there is evidence (Sloutsky, Lo, & Fisher, in press; Yamauchi & Markman, 2000) that adults are more likely than children to consider linguistic labels as category markers. There is also evidence that under different conditions adults may either rely on similarity for categorization (Smith & Sloman, 1994), or ignore it (Rips, 1989).

In short, the reported results support both quantitative and qualitative predictions of the model of label-as-attribute. As predicted, categorization appeared to be a function of two variables – the weight of linguistic label and the similarity ratios – and not of linguistic labels alone. These results also support the contention of the model that for young children linguistic labels are distinct attributes of entities. High correlations between the probabilities predicted by the model of similarity and the observed categorization frequencies support the hypothesis that categorization in young children is a similarity-based process.

Acknowledgments

This research has been supported by grants from the James S. McDonnell Foundation and the National Science Foundation to the first author.

References

Aslin, R., & Smith, L. (1988). Perceptual Development. *Annual Review of Psychology*, 39, 435-474.

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? *Journal of Experimental Child Psychology*, *64*, 3-26.

Birnholz, J. C., & Benaceraff, B. B. (1983). The development of human fetal hearing. *Science*, *222*, 516-518.

Estes, W. K. (1994). *Classification and cognition*. New York: Oxford University Press.

Jusczyk, P. W. (1998). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Lamberts, K., & Shanks, D. R. (Eds), (1997). *Knowledge, concepts and categories*. Cambridge, MA: MIT Press.

Mandler, J. M. (1997). Development of categorisation: Perceptual and conceptual categories. In G. Bremner, A. Slater, & G. Butterworth (Eds.), *Infant development: Recent advances* (pp. 163-189). Hove, England UK: Psychology Press.

Medin, D. (1975). A theory of context in discrimination learning. In G. Bower (Ed.), *The psychology of learning and motivation* (pp. 263-314), Vol. 9. New York: Academic Press.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Review of Psychology*, *35*, 113-138.

Napolitano, A., Sloutsky, V. M., & Boysen, S. (2001). *Proceedings of the XXIII Annual Conference of the Cognitive Science Society*.

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

Quinn, P., & Eimas, P. (1998). Evidence for global categorical representation of humans by young infants. *Journal of Experimental Child Psychology*, *69*, 151-174.

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

Roberts, K. (1995). Categorical responding in 15-month-olds: Influence of the noun-category bias and the covariation between visual fixation and auditory input. *Cognitive Development*, *10*, 21-41.

Roberts, K., & Jacob, M. (1991). Linguistic vs. attentional influences on nonlinguistic categorization in 15-months-old infants. *Cognitive Development*, *6*, 355-375.

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*, *6*, 1478-1492.

Sloutsky, V. M., & Lo, Y.-F. (2000). Linguistic labels and the development of inductive Inference. *Proceedings of the XXII Annual Conference of the Cognitive Science Society* (pp. 469-474). Mahwah, NJ: Erlbaum.

Sloutsky, V. M., Lo, Y.-F., & Fisher, A. V. (In Press). How much does a shared name make things similar? Linguistic labels and the development of inductive inference. *Child Development*.

Smith, E. E.& Sloman, S. A. (1994). Similarity- versus rule-based categorization. *Memory & Cognition*, 22, 377-386.

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