BUCLD 35 Proceedings
To be published in 2011 by Cascadilla Press

Effects of Labels on Children's Category Boundaries

Megan Johanson and Anna Papafragou

1. Introduction

Categorization, (defined as knowing when some things belong in a certain set to the exclusion of others) is generally recognized to be a complex process relying on a variety of cues, including perceptual (Imai, Gentner, & Uchida, 1994; Landau, Smith, & Jones, 1988; Quinn, Norris, Pakso, & Schmader, 1999; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002), conceptual (Booth & Waxman, 2002; Booth, Waxman, & Huang, 2005; Carmichael, Hogan, & Walter, 1932; Lupyan, 2007; Waxman & Namy, 1997) and linguistic information (Casasola & Bhagwat, 2007; Dessalegn & Landau, 2008; Rattermann & Gentner, 1998; Landau & Shipley, 2001; Plunkett, Hu, & Cohen, 2008). Here we investigate the role of language in categorization and propose that the influence of labels may not always be as privileged as previously thought.

Work with adults reveals that language appears to influence categorization. In a classic study by Carmichael and colleagues, adults were presented with a series of ambiguous images in conjunction with one of two labels. For example, the participants would view two small circles with a line attaching them. While viewing the image, half of the participants would hear "This figure resembles eyeglasses" and the other half would hear "This figure resembles dumbbells". After viewing the entire list, the subjects were asked to draw the images that they had seen. Participants did generally did not produce an accurate representation of the ambiguous image they saw. Rather, participants who heard "eyeglasses" were more likely to produce a representation resembling eyeglasses (e.g. an arched line connected the two circles) and those who heard "dumbbells" were more likely to produce a representation resembling dumbbells (e.g. a thicker line connected the two circles), demonstrating that labels are influential for ambiguous images. In addition, labels have been found to improve categorization even when they are completely redundant (Lupyan, Rakison, & McClelland, 2007).

---

* Both authors are affiliates of the University of Delaware. Contact Megan Johanson with any questions: johanson@udel.edu. This research was supported by NSF grant BCS#0641105 to Anna Papafragou. The authors would like to thank the members of the Language and Cognition Lab at UD as well as the Newark daycares.

Similarly, research with infants and children has revealed several situations where labels can affect categorization judgments. An experiment by Plunkett, Hu, and Cohen (2008) found that labels influence how infants perceive categories, but not when they are applied randomly. When the same label is applied to two different objects, it can boost the equivalence of two classes of objects (Landau & Shipley, 2001) and even lead children to make inductions about underlying characteristics of the objects (Sloutsky, Lo, & Fisher, 2001). Furthermore, the type of labels given in conjunction with an object can promote attention to different object properties (Waxman & Booth, 2001; Waxman & Braun, 2005; Yoshida & Smith, 2003). Finally, the presence of labels promotes the extraction of similarities between exemplars that underlie categorization (Casasola & Bhagwat, 2007). These experiments and others (Dessalegn & Landau, 2008; Rattermann & Gentner, 1998; Waxman & Booth, 2001) support the idea that language can influence categorization for infants and children

However, some research shows that language doesn't help all ages equally (Booth & Waxman, 2002; Balaban & Waxman, 1997; Sloutsky et al., 2001). We wanted to explore this for adults and 4 to 5 year old children. In particular, we want to know when labels help. Much of the previously mentioned work used extensive training with novel labels (Boothman et al., 2005; Casasola & Bhagwat, 2007; Smith et al., 2002). In contrast, we intend to explore the effect of language on categorization with only a single exposure to a cue.

Last, we are interested in whether other types of cues affect categorization to the same extent? Studies wilt non-label types of cues have revealed some interesting results. In a study with adults, Lupyan et al. (2007) found that location information (moving up or down to indicate where they live) did not help with categorization of "aliens" as much as labels (hearing a voice saying "leebish" or "gracious"). In addition, tones presented in conjunction with category exemplars do not help categorization for young children (Ferry, Hespos, & Waxman, 2010; Fulkerson & Waxman, 2007). Finally, demonstrating the function of an object can assist in categorization for children (Booth & Waxman, 2002). Expanding on this idea, we chose to explore the influence of perceptual facts as well as labels on categorization.

The goal for these experiments was to explore the influence of labels and perceptual facts on category boundaries for completely ambiguous objects. We chose to use novel labels ("is a blick/dax") and perceptual facts ("has a long/short beak"). We would expect that labels would improve participants' abilities to form categories based on all of the relevant evidence mentioned previously which demonstrates that labels are usually relevant cues to category membership (Casasola & Bhagwat, 2007; Landau & Shipley, 2001). In addition, we would expect that presenting participants with perceptual facts as cues would aid in category formation because they refer to physical facts about the objects that are clearly visible to compare and contrast between images. These experiments are the first to look specifically at how different types of cues can influence the categorization of perceptually ambiguous images for children. The results will reveal what types of cues children can take advantage of when there is no perceptual pattern to rely on.

2. Experiment 1

2.1. Method

2.1.1. Participants

Thirty-nine children and 39 adults participated. The children were between the ages of 4;1 and 5;6 with a mean age of 4;11. They were recruited from the Early Learning Center in Newark, DE. The adults were recruited from Introductory Psychology courses at the University of Delaware and received course credit for participating.

2.1.2. Materials

All participants saw four sets of images: two flower sets, a fish set, and a bird set. Each set had two standards and five target morphed images, which were 10, 30, 50, 70, and 90 percent like one of the standards (see Figure 1 for example set). For instance, in Figure 1, the bottom left image is 10% like standard 2 and 90% like standard 1. The images were morphed using FantaMorph, a commercial morphing program.

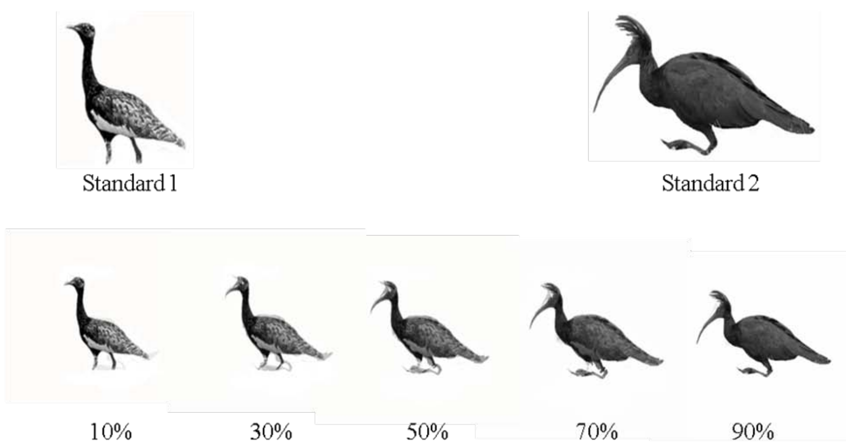

Figure 1. A sample set of morphed images.

Pictures corresponding to10%, 30%, 70%, and 90% were considered unambiguous images because they were more visually similar to one standard compared to the other. However, the 50% image was completely ambiguous between the two standards. A rating study with adults confirmed that the 50% images in all sets were in fact equally similar to each of the standards.

2.1.3. Procedure

Prior to the start of the experiment all participants were told that they were going to play a game where they would see different pictures and were instructed to try to remember what they saw because the experimenter would ask about the pictures later.

In the training phase, participants first viewed the two standards on the top half of the computer screen, and were given information about them, depending on the condition. Next, each target object was presented one at a time at the bottom center of the screen. The experimenter always gave information about the unambiguous targets that was consistent with its visual characteristics (it always grouped the target with the most perceptually similar standard). For each of the ambiguous targets, the experimenter informed half the participants that it "went with" standard 1 and half that it "went with" standard 2 for each set. The experimenter always pointed to the appropriate standard or target while giving information about it.

There were three between-subjects conditions: Label, Perceptual Fact, and No Cue. See Table 1 for the basic procedure layout. In the Label condition, participant heard different novel labels ("blick/dax") presented in conjunction with each of the standards. After the introduction of the target image, one of the two labels applied to the standards was repeated while pointing to the target. In the Perceptual Fact condition, the experimenter drew the participants' attention to different perceptual facts ("long/short beak") for each of the standards. After the introduction of the target image, one of the two perceptual facts that were applied to the standards was repeated while pointing to the target. Unlike the other two conditions, the experimenter provided no categorization information in the No Cue condition. Instead, the experimenter pointed to each image in turn while saying "Look at this one!" This condition was used at a baseline to see if the children had a bias to judge the 50% images as more similar to one standard than the other. Of interest was how the children would attend to and retain the information given in each condition. Would the presence of a certain type of cue lead to more accurate performance compared to the others?

At test, after viewing all five target images in a set, the participants viewed the same five target images in the same order. However, the two standards were switched to reduce some surface cues. Participants in all conditions heard: "Which one of these (pointing to each of the two Standards in turn), does this one go with?"

The experiment was designed in blocks with each set being a separate block. For half of the subjects the images within each block were in ascending order (10%, 30%, 50%, 70%, 90%) and for half they were in descending order (90%, 70%, 50%, 30%, 10%). At test, participants saw the target trials in the same order as before, but the position of the two standards was reversed in order to eliminate surface cues. The assignment of novel words/facts to standards was always consistent with the perceptual match for the unambiguous trials (10%, 30%, 70%, 90%) but was counterbalanced for the ambiguous trials (50%). The children completed the task in two sessions due to attention limitations.

Table 1. Presentation format for each condition in Experiment 1.

	Training			Test
	Standard 1	Standard 2	Target	Target
Label	"Look! This one is called a lorp!"	"Look! This one is called a pim!"	"Look! This one is a lorp/pim like this one!"	"Which one of these does this one go with?"
Percept. Fact	"Look! This one has a short beak!"	"Look! This one has a long beak!"	"Look! This one has a short/long beak like this one!"	"Which one of these does this one go with?"
No Cue	"Look at this one!"	"Look at this one!"	"Look at this one!"	"Which one of these does this one go with?"

2.2. Coding

For the Label and the Perceptual Fact conditions, correct responses were those that conformed to either the combination of perceptual similarity and the cue given (unambiguous trials) or cue alone (ambiguous trials). For the No Cue condition, correct responses were those that conformed to perceptual similarity for the unambiguous (10%, 30%, 70%, 90%) trials. For the ambiguous (50%) trials correct responses corresponded to Standard 1 half of the time and Standard 2 half of the time.

2.3. Results and Discussion
2.3.1. Unambiguous Trials

We first looked at the average proportion of correct responses on the unambiguous trials. That is, we collapsed the results from the 10%, 30%, 70%, and 90% trials. The results are at or near ceiling for both adults (M_{FACT} = .99, M_{LABEL} = 1.00, $M_{NO\ CUE}$ = .96) and children (M_{FACT} = .91, M_{LABEL} = .92, $M_{NO\ CUE}$ = .92), revealing that the participants understood the task and were able to form correct categories when labels and facts lined up with perceptual similarity.

2.3.2. Ambiguous Trials

Next we looked at the proportion of correct responses on the ambiguous (50%) trials. We ran an ANOVA looking at the proportion of correct responses as the dependent variable with Condition (Label, Perceptual Fact, No Cue) and Age Group (adults, children) as a between-subjects factors. The ANOVA revealed a significant effect of Age Group ($F(1, 5) = 24.27$, $p < .0001$; $M_{AD} = .79$; $M_{CH} = .54$) and Condition ($F(2, 5) = 7.91$, $p = .0008$). There was also a marginally significant Age Group by Condition interaction ($F(2,5) = 3.06$, $p = .053$). This interaction was caused by a significant difference between children and adults in both the Label ($F(1, 24) = 25.44$, $p < .0001$) and Perceptual Fact ($F(1, 24) = 12.83$, $p = .002$) conditions (see Figure 2). As expected, performance in the No Cue condition did not differ between adults and children ($F(1, 24) = .78$, $p = .39$, n.s.).

In addition, performance in the No Cue condition was not different from chance for either the children ($t(12) = -.28$, $p = .79$, n.s.) or the adults ($t(12) = .94$, $p = .37$, n.s.). This enables us to confidently say that the 50% images are perceptually ambiguous and to use the No Cue means as a baseline. Further t-tests show that adults' performance was significantly different from chance for both the Label ($t(12) = 17.73$, $p < .0001$) and Perceptual Fact ($t(12) = 7.68$, $p < .0001$) conditions. However, the children did not perform differently from chance for either the Label ($t(12) = 0.76$, $p = .46$) or Perceptual Fact ($t(12) = 1.81$, $p = .09$, n.s.) conditions.

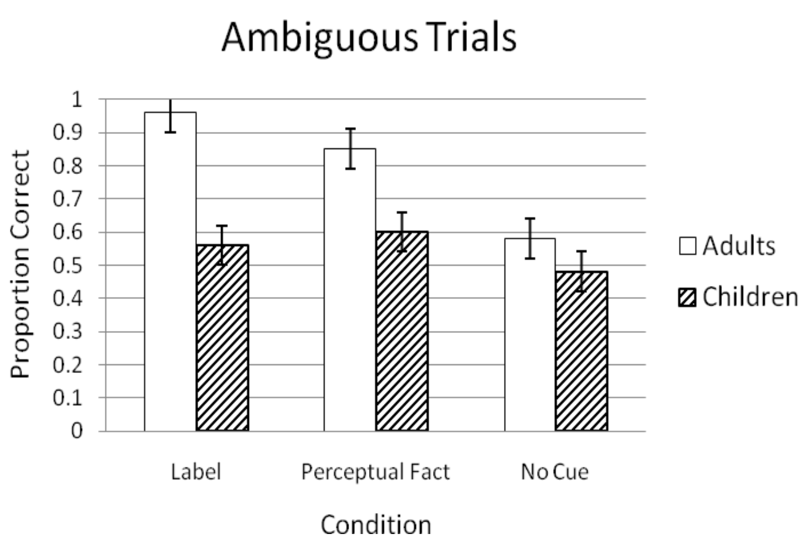

Figure 2. Proportion of correct responses on the ambiguous trials in Experiment 1.

Looking specifically at adults, we can see an effect of condition ($F(2, 36) = 12.39$, $p < .0001$). Performance in the Label condition is better than performance in the No Cue condition ($F(1, 24) = 20.00$, $p = .0002$). Perceptual Facts also lead to better performance than the No Cue condition ($F(1, 24) = 8.28$, $p = .008$). Interestingly, Labels were even better than Perceptual Facts at aiding categorization ($F(1, 24) = 4.91$, $p = .04$). For children alone, there was no effect of condition ($F(2, 36) = 0.75$, $p = .49$, n.s.), which is very surprising. Children did not take advantage of perceptual facts or labels to aid categorization.

In sum, our results reveal that for adults, both labels and perceptual facts can sway categorization of perceptually ambiguous stimuli. In fact, labels are even better than perceptual facts, presumably because labels are better indicators of category membership than perceptual facts. In contrast, neither labels nor perceptual facts altered performance on ambiguous stimuli for children. This should be considered surprising due to the previous work demonstrating that labels help children form categories (Casasola & Bhagwat, 2007; Dessalegn & Landau, 2008; Rattermann & Gentner, 1998; Landau & Shipley, 2001; Plunkett et al., 2008).

These differences between children and adults could be the result of a few things. First, children may forget the labels and perceptual facts given to them during training. Each test trial occurred five trials after corresponding training trial. The delay in categorization might have been too long for the children to retain and implement the cues. Furthermore, training is brief. Participants received only a single exposure to each new label/perceptual fact. The single exposure training might have been insufficient for children to form a reliable connection between the cues and the images.

Second, children, unlike adults, may be unable to use internal cues for category structure given by trial order within a set. Recall that the presentation order within a set was either ascending or descending (i.e., from 10% to 90% or vice versa). This means that the same label/perceptual fact is given for either the first 2 or 3 trials (depending on whether the 50% image corresponds to standard 1 or 2) before switching to the other standard's label/perceptual fact. In theory, the adults could only pay attention to the trial in which the cue changed but not fully integrate the cue and image. This could have resulted in inflated success rates for adults, but not for the children.

In an effort to reduce these potential confounds of the first experiment, we performed a second experiment with a few changes. Each test trial was presented immediately after corresponding training trial to minimize demands on memory. Furthermore, trials were randomized within each set to remove internal cues about category structure.

In addition, we wanted to explore the difference in adult performance between the Label and Perceptual Fact condition. Adults performed significantly better in the label condition, even though one could hypothesize that perceptual facts may be reinforcing visible features and therefore placing fewer demands on memory than the label condition (where adults had nothing to help them perform the task except for remembering the novel label). To understand these results better, we added an Arbitrary Fact condition. This condition allows us to explore whether providing a fact that does not refer to physical features of the objects would result in performance similar to that in the Label or Perceptual Fact condition. Our prediction would then be that arbitrary facts (e.g., “This one drinks milk”) should pattern with labels, not perceptual facts because they are not easily traced back to the image like the perceptual facts are (“This one has a long beak”).

3. Experiment 2
3.1. Method
3.1.1. Participants

Sixty-three children and 48 adults participated. The children were between the ages of 4;0 and 5;3 with a mean age of 4;6. They were recruited from several daycares in Newark, DE including the Early Learning Center, Newark Methodist Preschool, Newark Christian Childcare, and Panda Early Education Center. The adults were recruited from Introductory Psychology courses at the University of Delaware and received course credit for participating.

3.1.2. Materials

The materials were the same as in Experiment 1.

3.1.3. Procedure

In the training phase, participants saw the same screen as in the training phase of Experiment 1. However, after hearing the cue, the experimenter would press a key and the target would disappear from the screen while the two standards would remained (See Table 2). The experimenter would say, “Oops! It went away!” Following another key press, the target would reappear in the same position. The experimenter would then say, “Look! Here it is again! Which one of these does it go with?” This process would be repeated for each target object, grouped in blocks (sets). Within each block (set) targets would appear in random order.

Two presentation orders were made with the Target Objects appearing in random order and by counterbalancing the position of the standard objects. For List 1, Standard 1 was on the left but for List 2, Standard 2 was on the left. Once those lists were complete we created a reverse list for each, resulting in a total of four presentation orders.

Table 1. Presentation format for each condition in Experiment 2.

Standard 1	Standard 2	Target Appears	Target Disappears	Target Reappears

Label	"Look! This one is called a lorp!"	"Look! This one is called a pim!"	"Look! This is another lorp/pim!"		"Which one of these does this one go with?"
Fact	"Look! This one has a short beak!"	"Look! This one has a long beak!"	"Look! This is another one with a short/long beak!"		"Which one of these does this one go with?"
Arbitrary Fact	"Look! This one drinks milk!"	"Look! This one drinks water!"	"Look! This is another one that drinks milk/ water!"		"Which one of these does this one go with?"
No Cue	"Look at this one!"	"Look at this one!"	"Look at this one!"		"Which one of these does this one go with?"

3.2. Coding

The coding was the same as in Experiment 1.

3.3. Results and Discussion

3.3.1. Unambiguous Trials

We began by looking at the average proportion of correct responses on the unambiguous trials. The results are all at or near ceiling for both children (M_{FACT} = .98, M_{LABEL} = .96, M_{AF} = .95, $M_{NO\ CUE}$ = .96) and adults (M_{FACT} = 1.00, M_{LABEL} = 1.00, M_{AF} = .99, $M_{NO\ CUE}$ = .98), revealing that the participants in all conditions understood the task and were able to form correct categories when labels and facts lined up with perceptual similarity.

3.3.2. Ambiguous Trials

Next we looked at the proportion of correct responses on the ambiguous (50%) trials. We ran an ANOVA looking at the proportion of correct responses as the dependent variable with Condition (Label, Perceptual Fact, Arbitrary

Fact, No Cue) and Age Group (adults, children) as a between-subjects factors. The ANOVA revealed a significant effect of Age Group ($F(1, 7) = 5.73$, $p = .02$; $M_{AD} = .81$; $M_{CH} = .71$) and Condition ($F(3, 7) = 20.73$, $p < .0001$). There was also a marginal Age Group by Condition interaction ($F(3, 7) = 2.73$, $p = .05$). This interaction was caused by a significant difference between children and adults in the Label ($F(1, 26) = 10.22$, $p = .004$) condition (see Figure 3). However, there was no difference in performance between children and adults in the Perceptual Fact ($F(1, 26) = 0.82$, $p = .37$) or Arbitrary Fact ($F(1, 26) = 2.26$, $p = .14$, n.s.) conditions. Again, performance in the No Cue condition did not differ between adults and children ($F(1, 25) = .44$, $p = .51$, n.s.). This means that children performing like adults in all conditions except for the Label condition.

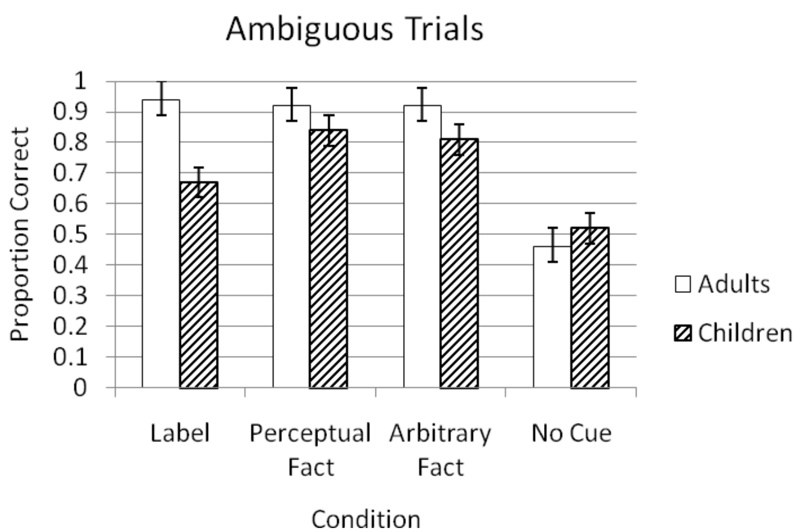

Figure 3. Proportion of correct responses on the ambiguous trials in Experiment 2.

Furthermore, performance in the No Cue condition did not differ from chance: adults ($t(11) = -0.62$, $p = .55$, n.s.), children ($t(14) = .29$, $p = 0.77$, n.s.). This means that our ambiguous images truly were ambiguous for both age groups. Additional t-tests show that the adults performed differently than chance in the Label ($t(12) = 13.40$, $p < .0001$), Perceptual Fact ($t(12) = 6.50$, $p < .0001$) and Arbitrary Fact ($t(12) = 8.86$, $p < .0001$) conditions. The children also performed differently from chance in the Label ($t(15) = 2.55$, $p = .02$), Perceptual Fact ($t(15) = 6.82$, $p < .0001$) and Arbitrary Fact ($t(15) = 6.45$, $p < .0001$) conditions. Thus, adults and children are able to use all three types of cues to adjust category boundaries.

To confirm the influence of the cues for adults we ran pair wise comparisons. All conditions compared to the No Cue condition were significantly different (p's < .0001). All other comparisons were not significant. This means that adults are able to use labels, perceptual facts, and arbitrary facts as cues to category membership, even when the image is visually ambiguous between the categories.

We performed the same pair wise comparisons for children alone and found that there was a significant difference between performance in the No Cue and Perceptual Fact ($F(1, 29) = 18.58$, $p = .0002$) and the Arbitrary Fact ($F(1, 29) = 15.77$, $p = .0004$) conditions . Surprisingly, there was also a marginal difference in performance between the Label and Perceptual Fact ($F(1, 30) = 4.17$, $p = .05$) and No Cue ($F(1, 29) = 3.05$, $p = .09$) conditions. However, there was no difference between performance in the Label and Arbitrary Fact ($F(1, 30) = 2.87$, $p = .10$, n.s.) conditions. Finally, we found no difference between the Perceptual and Arbitrary Fact ($F(1, 30) = 0.20$, $p = .66$, n.s.) conditions. In sum, we found that for children, labels result in performance differently from chance, but only marginally different from the No Cue condition. Both the Arbitrary and Perceptual Fact conditions appear to be slightly more privileged than the Label condition for children, which is the opposite of the adults' performance in the first experiment.

4. General Discussion

For adults, labels, perceptual facts, and arbitrary facts help adjust category boundaries. In general, labels are very reliable indicators of category membership (Balaban & Waxman, 1997; Fulkerson & Waxman, 2007), as seen in Experiment 1. The task demands of Experiment 1 did not allow for children to use labels, perceptual facts, or arbitrary facts as cues to adjust category boundaries for ambiguous stimuli. However, the modified demands of the Experiment 2 allowed children take advantage of the all the cues resulting in performance different from chance in all cue conditions.

There are some differences that remain from the first to the second experiment, specifically between the adults and children in the Label condition. While children are performing differently from chance, and marginally better than the No Cue condition, they are still not using labels to the same extent as the adults are.

Also, contrary to our predictions and the performance of adults in Experiment 1, labels appear to be less privileged than either Perceptual or Arbitrary labels. One might claim that the limited exposure to the cues might have been the reason why children could not use any cues in Experiment 1 and why their use of labels in Experiment 2 was not adult-like. However, children's performance in the Arbitrary Fact condition would argue against this. If children are able to remember an arbitrary piece of information in the form of the fact, they should also be able to remember an arbitrary label. Since they are performing similarly to adults in the Arbitrary Fact condition, this cannot be the best explanation.

This difference could reveal how labels influence children's categorization at the subordinate level (as in 3 of the 4 sets our task) as opposed to the basic level, where most of the previous research has focused (Sloutsky et al., 2001; Imai et al., 1994; Lupyan, 2008; Mervis, Golinkoff, & Bertrand, 1994; Waxman & Namy, 1997). Little work has looked at children's experiences with subordinate level categories, and the small amount that has been done only looked at children's preferred taxonomic level of labeling (Blewitt, Golinkoff, & Alioto, 2000). Extensions of this work could reveal the specific taxonomic level(s) at which children are able to use certain types of information to make decisions about categorization in adult-like ways. That is, labels might have a stronger influence for children on categorization within a basic level category and less so for subordinate level categories. Adults, however, may have no problem conforming to the task demands and accepting the novel labels as subordinate object labels.

References

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

Blewitt, Pamela, Golinkoff, Roberta, & Alioto, Anthony. (2000). Do toddlers have label preferences? A possible explanation for word refusals. *First Language, 20*, 253-272.

Booth, Amy & Waxman, Sandra. (2002). Word learning is 'smart': Evidence that conceptual information affects preschoolers' extension of novel words. *Cognition, 84*, B11-B22.

Booth, Amy, Waxman, Sandra, & Huang, Yi. (2005). Conceptual information permeates word learning in infancy.

Carmichael, L, Hogan, HP, & Walter, AA. (1932). An experimental study of the effect of language on the reproduction of visually perceived form. *Journal of Experimental Psychology, 15*, 73-86.

Casasola, Marianella & Bhagwat, Jui. (2007). Do novel words facilitate 18-month-olds's spatial categorization? *Child Development*, *78*, 1818-1829.

Dessalegn, Banchiamlack & Landau, Barbara. (2008). More than meets the eye: The role of language in binding and maintaining feature conjunctions. *Psychological Science, 12*, 189-195.

Ferry, Alissa, Hespos, Susan, & Waxman, Sandra. (2010). Categorization in 3- and 4-month-old infants: An advantage of words over tones. *Child Development, 81*, 472-479.

Fulkerson, Anne & Waxman, Sandra. (2007). Words (but not Tones) Facilitate Object Categorization: Evidence From 6- and 12-Month-Olds. *Cognition. 105*, 218-228.

Imai, Mutsami, Gentner, Dedre, & Uchida, Nobuko. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition, *Cognitive Development, 9*, 45-75.

Landau, Barbara & Shipley, Elizabeth. (2001). Labelling patterns and object naming. *Developmental Psychology, 4*, 109-118.

Landau, Barbara, Smith, Linda, & Jones, Susan. (1988). Object perception and object naming in early development. *Trends in Cognitive Sciences*, *2*, 19-24.

Lupyan, Gary, Rakison, David, & McClelland, James. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science, 18*, 1077-1083.

Lupyan, Gary. (2007). Reuniting categories, language, and perception. Presented at *The 29th Annual Conference of the Cognitive Science Society*. Nashville, TN.

Lupyan, Gary. (2008). From chair to "chair": A representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General, 137*, 348-369.

Mervis, Carolyn, Golinkoff, Roberta, & Bertrand, Jacquelyn. (1994). Two-year-olds readily learn multiple labels for the same basic-level category. *Child Development, 65*, 1163-1177.

Plunkett, Kim, Hu, Jon-Fan, & Cohen, Leslie. (2008). Labels can override perceptual categories in early infancy. *Cognition, 106*, 665-681.

Quinn, Paul, Norris, Christopher, Pakso, Rita, & Schmader, Tayna. (1999). Formation of a categorical representation for the spatial relation of between by 6- to 7-month-old infants. *Visual Cognition, 6*, 569-585.

Ratterman, Mary & Gentner, Dedre. (1998). The effect of language on similarity: The use of relational labels improves young childrens' performance in a mapping task. In K. Holyoak, D. Gentner, and B. Kokinov (Eds.), *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences.* (pp. 274-282). Sophia: New Bulgarian University.

Sloutsky, Vladimir, Lo, Ya-Fen, & Fisher, Anna. (2001). How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child Development, 72,* 1695-1709.

Smith, Linda B, Jones, Susan S, Landau, Barbara, Gershkoff-Stowe, Lisa, & Samuelson, Larissa. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13*, 13-19.

Waxman, Sandra & Booth, Amy. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology, 43*, 217-242.

Waxman, Sandra & Braun, Irena. (2005). Consistent (but not variable) names as invitations to form object categories: New evidence from 12-month-old infants. *Cognition, 95*, B59-N68.

Waxman, Sandra & Namy, Laura. (1997). Challenging the notion of a thematic preference in young children. *Developmental Psychology, 33*, 555-567.

Yoshida, Hanako & Smith, Linda. (2003). Known and novel noun extensions: Attention at two levels of abstraction. *Child Development, 74*, 564-577.