Cognitive Development, 14, 487–513 (1999)
ISSN 0885-2014

Comparison in the Development of Categories

Dedre Gentner
Northwestern University, Evanston, IL, USA

Laura L. Namy
Emory University, Atlanta, GA, USA

Recent research on children's word learning has led to a paradox. Although word learning appears to be a deep source of insight into conceptual knowledge for children, preschoolers often categorize objects on the basis of shallow perceptual features such as shape. The current studies seek to resolve this discrepancy. We suggest that comparing multiple instances of a category enables children to extract deeper relational commonalities among category members. We examine 4-year-olds' categorization behaviors when asked to select a match for a target object (e.g., an apple) between a perceptually similar, out-of-kind object (e.g., a balloon) and a perceptually different category match (e.g., a banana). Children who learn a novel word as a label for multiple instances of the category are more likely to select the category match over the perceptual match. Children who learn a label for only one instance are equally likely to select either alternative. This effect is present even when individual target instances are more perceptually similar to the perceptual choice than to the category choice. We conclude that structural alignment processes may be important in the development of category understanding.

How do children learn conceptual structure? Recent research has demonstrated that even very young children have some insight into the nature of categories (Gelman & Markman, 1986; Mandler & Bauer, 1988; Mandler, Bauer, & McDonaugh, 1991; Waxman & Markow, 1995). Nonetheless, there is evidence of substantial gains in children's understanding of categories and concepts during infancy and the preschool years (Mandler, 1992; Nelson, 1973; Vygotsky, 1962),

This research was supported in part by the National Science Foundation grant SBR-9511757 awarded to the first author. We thank Sandra Waxman and Phillip Wolff for comments on this paper and Linda Smith for discussions of this research. We also thank Sheila Singh, Marjorie Jones, and Michelle Osmundson for help in data collection. Finally, we thank two anonymous reviewers for their suggestions.

Direct all correspondence to: Dedre Gentner, Department of Psychology, Northwestern University, 2029 Sheridan Rd., Evanston, IL 60208; E-mail: gentner@nwu.edu

Manuscript received January 12, 1998; revision accepted June 16, 1999

as evidenced by changes in patterns of habituation (Balaban & Waxman, 1997; Madole, Oakes, & Cohen, 1993; Oakes, Plumert, Lansink, & Merryman, 1996; Waxman & Markow, 1995); sequential touching and sorting behaviors (Mandler & Bauer, 1988; Nelson, 1973; Ricciuti, 1965; Sugarman, 1983); and spontaneous naming behavior (Clark, 1973; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976).

Many researchers have focused on the word-extension task as a way of exploring children's understanding of categories. Children are taught a new word that is exemplified with one instance, and then are asked to extend the word to other potential exemplars. Extensive research by Markman, Waxman, and others has demonstrated that, whereas preschool children may group objects together for a wide variety of different reasons, including thematic relatedness, when they are taught a new word, they seem to assume strongly that the word applies to things of like kind (Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Imai, Gentner, & Uchida, 1994; Markman, 1989; Markman & Hutchinson, 1984; Waxman & Gelman, 1986; Waxman & Hall, 1993; Waxman & Markow, 1995). This propensity of children to extend word meanings according to like kinds has attracted a great deal of research interest, as it seems to offer a window into a mechanism by which children may acquire command of the categories that characterize their world.

For example, S. Gelman and colleagues (Gelman & Coley, 1990; Gelman & Markman, 1986; Gelman & Wellman, 1991; Kalish & Gelman, 1992) have demonstrated that preschool children who are taught a novel non-obvious property or characteristic of a standard object will extend that property to other instances of that kind when the standard object is named. In addition, Kemler Nelson (1995) has found that preschool-aged children who are shown an object that clearly affords a particular function (e.g., a paintbrush) and hear a novel name for this object extend the novel name to other objects that afforded the same function as the original object.

These findings suggest that young children appreciate that words refer to categories that are organized around deep, core properties. Here, however, the evidence is somewhat equivocal. On the one hand, there is considerable evidence consistent with the claim that children's early word learning invites deep categories. For example, in a triad selection task, young children shown a standard object are more likely to choose a categorically related object over a thematic match if the standard is labeled—"This [standard] is a dax. Can you find another dax?"—than if asked to "find another one" without labeling the standard (Markman, 1989; Markman & Hutchinson, 1984; Waxman & Gelman, 1986; Waxman & Kosowski, 1990). Markman's *taxonomic constraint* and Waxman's *noun-category linkage* are principles embodying the idea that young children implicitly know that the meanings of nouns are organized around categorical relations rather than thematic associations (Markman, 1989; Waxman, 1990).

Yet, despite this evidence that children acquire a deeper appreciation of category structure in the context of naming, there is also considerable evidence that, especially early in development, children often seem to think of "like kinds" as "perceptually like kinds," especially "shape-similar kinds" (Baldwin, 1989, 1992; Gentner, 1978; Gentner & Imai, 1995; Golinkoff, Shuff-Bailey, Olguin, & Ruan, 1995; Imai et al., 1994; Landau, Smith, & Jones, 1988; Smith, Jones, & Landau, 1992; Ward et al., 1989).

Imai et al. (1994) explored this early focus on perceptual over conceptual features of categories by giving 3-year-olds and 5-year-olds a word-extension task in which taxonomic relatedness was separated from shape-similarity. Children were taught a new word in "dinosaur language," such as "dax" for an apple, and then were asked to choose the other "dax" from among a banana (category choice), a tennis ball (shape choice), and an apple tree (thematic choice). The results confirmed the classic finding that children in the word task shift away from the thematic choice, relative to children given a non-word task of the form "See this? Which one does it go with?" However, the shift was not to the category choice but rather to the shape-similar choice. In fact, 3-year-olds were significantly *less* likely to choose the category alternative in the word task than in the non-word task. Even 5-year-olds showed a preference for the same-shape item in a word-extension task (although they were significantly more likely to make a category choice in this task than were the 3-year-olds). Imai et al. (1994) characterized this development as a "shape-to-category-shift." Golinkoff et al. (1995), also found a shift from a reliance on shape and other perceptual commonalties in 4-year-olds to more adult-like word extensions at seven years of age. Further, as discussed below, Gentner and Imai (1995) found that even when children were given the option of selecting a high-similarity category choice (such as a pear), they were equally likely to choose a high-similarity foil in the novel word-extension task.

Baldwin (1992) administered a similar task in which she gave 3- to 5-year-olds a choice between a thematic alternative and either a same-shape or a same-category alternative, and found that children shifted from thematic to either shape or category, responding in a word extension task. However, when offered a choice between a same-shape alternative and a same-category alternative, children in her study—as in the Imai et al. (1994) study—showed a significant shift towards shape responding (and away from category responding) in the word-extension task. Carrying the question one step further, Landau, Smith, and Jones (1998) tested the possibility that children would extend words on the basis of function when the function of objects was clearly demonstrated. However, they found—strikingly—that 2-year-olds continued to extend word meanings on the basis of shape even when a cross-cutting function, such as that of a sponge (soaking up water), was demonstrated and objects sharing that function—but not sharing shape—with the standard object were available. Children in this situation extended on the basis of shape rather than function (see also Gentner, 1978; Landau et al., 1988).

These findings would seem to suggest that word learning could actually be misdirecting children by inviting a focus on shallow, perceptual commonalities. How can we reconcile the evidence that children appear to extend words on the basis of a rather shallow perceptual heuristic with the evidence that language learning is a source of deep insight for children?

One possibility, based on research in analogical learning, is that even if initial extensions are perceptually based, later comparisons among instances may give rise to deeper insights into category structure. Such comparisons could be prompted by common perceptual features, by a common label, or by both. The key to this proposal is modeling comparison as a structure-mapping process that—even if initially prompted by common perceptual features—acts to render common relational structure more salient. Clearly, such a mechanism would fill an important role in children's categorization, because the core, theorylike aspects of concepts are often inherently relational: for example, causality relations, progeneration relations, and predator-prey relations.

This proposal differs from traditional accounts of similarity-based abstraction. The traditional featural account runs roughly as follows: as instances of a category are compared, common features are retained and distinctive features are discarded, resulting in a representation that is more abstract in the sense of having fewer features than the original exemplars. For example, comparing two exemplars—*red/ square/ large* and *blue/square/large*—would result in the recognition that color is irrelevant, resulting in the category representation *square/large.* Such abstractions are essentially subtractive. The removal of distinctive concrete features leads to a more general representation, which will then be compatible with more new examples. Thus in classical transfer studies, the greater the breadth of training, the greater the breadth of transfer. For example, Ross, Nelson, Wetstone, and Tanouye (1986) found that poor learners could generalize to more new instances of a concept if they had seen three instances of the concept during training than if they had seen only one.

Although our proposed mechanism also promotes commonalities among exemplars, it does so by a radically different mechanism. The process of structural alignment is one that operates to promote common relational structure, because of its sensitivity to relational connections and preference for deep common systems, as explained below. Thus when two representations are aligned, common structure is preferentially highlighted. This aspect of structural alignment is important for children's learning, we suggest, because it acts to elevate the salience of relational knowledge that might otherwise remain shadowy and implicit. There is considerable evidence that comparison leads to structural alignment (Gentner & Markman, 1997; Gentner, Rattermann, & Forbus, 1993; Gick & Holyoak, 1983; Goldstone, Medin, & Gentner, 1991; Markman & Gentner, 1993a, 1993b, 1996; Medin, Goldstone, & Gentner, 1993). For example, Markman and Gentner (1993b) used a simple mapping task with adults to demonstrate that carrying out a comparison promotes relational alignment. Participants were shown two scenes. Scene 1 shows a truck towing a car; scene 2 shows a (similar) car towing a boat. When asked which object in scene 2 best corresponded to the car in scene 1, the majority (not surprisingly) chose the highly similar car. But when a second group of participants was first asked to rate the similarity of the two scenes, and then given the same mapping task, there was a dramatic change. These subjects selected the boat in scene 2 as the best match for the car in scene 1—that is, they chose the relational match (tow-ee → tow-ee instead of car → car). We infer that the act of carrying out a similarity comparison induced a structural alignment in which the common relational structure became sufficiently salient to induce correspondences based on structural roles:

1. **tows** [car, boat]
2. **tows** [truck, car].

What these results suggest is that the very act of carrying out a comparison promotes structural alignment and renders common relational structure more salient. Thus, comparison can serve to highlight previously implicit relational commonalities, such as common causal structure.

There is considerable support for this claim from the research on analogical learning, in which deriving common relational structure is a central goal (Falkenhainer, Forbus, & Gentner, 1989; Gentner, 1983; Gentner & Markman, 1997; Holyoak & Thagard, 1989; Hummel & Holyoak, 1997; Keane & Brayshaw, 1988). According to structure-mapping theory, the process of comparison begins with local matches—such as object and part matches—which coalesce during processing to yield a maximal structurally consistent[1] alignment (for details see Falkenhainer et al., 1989; Forbus, Gentner, & Law, 1995). An important constraint on the structural alignment process is *systematicity*: A predicate that is relationally connected to other mappable predicates—for example, by higher-order causal relations—is more likely to be mapped than an isolated predicate. Systematicity acts to promote interconnected relational systems over isolated attributes.

The beneficial effects of comparison hold for mundane similarity as well as distant analogies. For example, Kotovsky and Gentner (1996) found that preschool children who compared perceptually similar pairs of pictures were subsequently better able to notice cross-dimensional correspondences based on the same matching relations (symmetry and monotonic increase) than were control children. Even without instruction or feedback, experience with same-dimension pairs facilitated the matching of cross-dimension pairs. Namy, Smith, and Gersh-

[1] A *structurally consistent* alignment is one that maintains a one-to-one mapping (i.e., an element in one representation corresponds to at most one element in the other representation) and parallel connectivity (i.e., if predicates correspond across the two representations, their arguments must correspond as well).

koff-Stowe (1997) demonstrated that 18-month-old infants who compared instances of two object kinds in a sorting task were more likely to later spontaneously sort the objects into two categories than those who did not experience comparison. Loewenstein and Gentner (1999; Loewenstein, 1997) found that children performed better in a mapping task from one model room to another (Deloache, 1995) when they had previously compared two nearly identical instances of the initial room. There is evidence, then, that the act of comparing perceptually similar entities can result in the extraction of deeper structural commonalities.

We suggest that comparison mechanisms may provide a way to resolve the conflict between the evidence that children use shallow perceptual cues in word learning and the equally compelling evidence that word learning is a deep source of insight for children. Based on the above line of reasoning, reliance on surface-level perceptual information could in some cases even be *constructive* in the following way. First, suppose that comparison can be invited by perceptual similarity and/or by common labels. Suppose further that when the items share deeper relational commonalities, the comparison process makes this common structure more apparent. When the items share *both* perceptual and relational commonalities—as is true for basic level categories, for example—an initial focus on perceptual similarities could lead the child to comparisons that yield deeper commonalities.

There is a hint that this kind of process may be operating in children's word extensions in a study conducted by Gentner and Imai (1995). This study was an extension of the Imai et al. (1994) study described above, in which children were taught a new word (e.g., *blicket*) for a pictured object (e.g., an apple) and asked to find another *blicket.* Children (especially 3-year-olds) overwhelmingly selected a balloon (shape match) over a banana (category match) as the *blicket.* These results suggest a strong perceptual bias. However, it also seemed possible that children might understand the conceptual nature of word meanings, but still expect category members to be perceptually similar. In this case, they might simply have been confused by the separation of perceptual and conceptual similarity in this task. Therefore, Gentner and Imai designed a more sensitive test of category knowledge by adding an alternative—for example, a pear—that was *both* shape-similar and categorically related to the standard. For example, given an apple (the *blicket*) as standard, children were asked to select another blicket from among four pictured alternatives:

Banana	Balloon	Pear	Tree
(category)	(shape)	(category+shape)	(thematic)

After the first selection, the selected item was placed next to the standard and the child was asked to choose another blicket. The results were surprising in two respects. First, the initial selections were based on shape: 80% of the children selected either the balloon or the pear. Further, responses were evenly divided between these two same-shape choices, despite the fact that the pear also shared category membership with the standard. The second interesting feature was the second-round behavior. Not surprisingly, the children who had selected the shape-only alternative (the balloon) on the first round went on to choose the other same-shape match, the pear. But the reverse behavior did not occur. Children who had selected the shape + category match (the pear) on the first round tended to choose the category match (the banana) on the second round, despite the availability of a same-shape alternative (the balloon). Why? One rather mundane explanation is that this second-round improvement simply reflects that children who selected the pear on the first round had superior initial category knowledge. However, an intriguing possibility is that the juxtaposition of the standard with the first selection (apple with pear) invited an alignment that made other commonalities salient, such as the functional commonality of being edible.

The goal of the present study is to examine directly the possibility that explicit comparison of perceptually and conceptually similar objects can yield insight into conceptual structure in preschool-aged children. In this study we tested the possibility that comparing multiple instances of a category may lead to deeper conceptual insights. It seems likely that children's acquisition of category knowledge in everyday life is facilitated by exposure to multiple instances of a category. This hypothesis is consistent with the literature on adult category learning (Elio & Anderson, 1984; Medin & Ross, 1989; Skorstad, Gentner, & Medin, 1988), indicating that exposure to multiple instances of a category can lead to category abstraction.

We tested this hypothesis by administering a forced-choice word extension task similar to the ones described above. We experimentally manipulated the children's opportunity to compare objects from a given category by presenting children with either a single standard instance or multiple instances of the category before eliciting their word extensions. If comparing multiple instances of a category yields insight into deeper commonalities, then children viewing multiple instances should be more likely to select a match on the basis of taxonomic commonalities as opposed to perceptual similarities, whereas children viewing a single instance should focus predominantly on perceptual similarities.

EXPERIMENT 1

In this experiment, we seek evidence that comparing multiple instances of an object category enables preschool-aged children to abstract conceptual or structural commonalities among instances of the category that are not readily evident from the perceptual input. To examine this hypothesis, we administer a forced-choice categorization task in which the children are asked to select between two alternatives as a match for the standard's category. In each set, one of the two alternatives, the category choice, is from the same category as the standard but is perceptually quite different from the standard. The other alternative, the perceptual choice, is a member of a different object category but perceptually resembles the standard. If children determine a match on the basis of perceptual commonalities, they should select the perceptual alternative over the category alternative when given a standard and asked to extend its label. Assuming this is the case, then the key question is whether being induced to compare perceptually similar items will lead children to extract a deeper conceptual structure. If so, in this case they should select the category alternative over the perceptual alternative.

In this study, children either viewed a single instance of the target category or several instances before making a choice. In the *no-compare* control condition, children saw only one member of the category (i.e., a single standard) before making a match. In the *compare* condition, children saw four discriminably different standards, all members of the target category, before making a choice. The stimuli were designed and piloted to achieve a level of moderate perceptual salience. Our aim was to ensure that the perceptual match for any single target would be attractive but not overpowering, so that we could assess the influence of our manipulation on children's response patterns.

The prediction is that children in the *no-compare* condition will select a match largely on the basis of perceptual features such as shape, orientation, and distinctive features, and will therefore be more likely to select the perceptual alternative than the category alternative. However, in the *compare* condition, children will have the opportunity to compare objects within the target category, and so may abstract deeper relational commonalities among the objects. Thus, we predict that those in the *compare* condition will select the category alternative more frequently than those in the *no-compare* condition.

Method

Participants. Sixteen 4-year-olds (mean age = 4;2, range = 4;0–4;8) participated. The children were from predominantly white upper-middle-class families in the greater Chicago area; their parents had responded to direct mailings or newspaper advertisements.

Materials. Forty-eight colored line drawings of real objects were used as stimuli. These drawings were organized into eight sets of six cards each. Each stimulus set included four standards and two choice alternatives. The four standards were discriminably different members of a superordinate-level object category (e.g., an apple, a pear, a slice of watermelon, and a bunch of grapes from the fruit category). The categories selected represent a range of taxonomic levels. The two choice alternatives were designed such that one card from each set, the perceptual match, was outside of the taxonomic category but perceptually resembled the standards (e.g., a balloon). The other alternative, the category match, was from the same category as the standards, but was perceptually distinct from the standards (e.g., a banana). Color of the pictures was controlled for in each set, such that no standard matched either choice card in color. A complete list of stimuli can be found in Table 1. Figure 1 displays a sample set.

Table 1. Materials Used in Experiment 1

				Alternatives	
Standard 1	**Standard 2**	**Standard 3**	**Standard 4**	**Perceptual**	**Taxonomic**
1. Apple	Pear	Watermelon	Grapes	Balloon	Banana
2. Plate	Bowl	Platter	Cake Pan	Cookie	Casserole Dish
3. Drum	Tambourine	Banjo	Trumpet	Hat Box	Flute
4. Carrot	Corn	Bell Pepper	Broccoli	Rocket	Turnip
5. Ice Cream	Lollipop	Candy Cane	Taffy	Top	Candy Bar
6. Cap	Derby Hat	Police Hat	Straw Hat	Igloo	Sombrero
7. Bicycle	Tricycle	Scooter	Roller Blade	Glasses	Skateboard
8. Caterpillar	Snake	Lizard	Frog	Rope	Turtle

Stimulus selection. To ensure proper control of the similarity structure, similarity ratings were obtained from a group of 48 Northwestern University psychology undergraduates. We asked the students to evaluate the perceptual similarity of the two alternatives to each of the four standards. Each student viewed only one standard from each set. We instructed students to judge "how similar-looking" each of the alternatives was to the standard, using a Likert scale on which a rating of 7 indicated "extremely similar," a rating of 4 denoted "somewhat similar," and a rating of 1 denoted "not at all similar." Subjects were told that the stimuli would be used in a study with young children, and were explicitly instructed to judge only the appearance of the objects.

As was intended, the four standards varied in their similarity to the perceptual and the category choices. However, overall ratings confirmed that the standards were more perceptually similar to the perceptual alternative ($M = 3.70$) than to the category alternative ($M = 3.24$), $F(3, 44) = 27.72$, $p < .001$.

Procedure. The child was seated at a low table in a quiet room in the laboratory, across from the experimenter. Parents were seated behind the children and were instructed not to speak to the children during testing.

The children were randomly assigned to either the *compare* or the *no-compare* condition. In both conditions, children were introduced to "Jo-Jo," a toy dog, and were told that they would learn Jo-Jo's special names for things. In both conditions, the experimenter labeled the standard category with a novel noun and then asked the children to help Jo-Jo find another object that bore the same label.

In the *no-compare* control condition, the experimenter showed the child a *single* standard and labeled it with a novel noun, for example, "This is a blicket." The standard in each set that was rated most perceptually similar to the perceptual alternative by the adults was used as the single standard card for all subjects. After labeling the object, the experimenter asked the child to repeat the novel word. She then laid the two alternatives on the table (see Fig. 1a) and asked the child, "Which one of these is a blicket?"

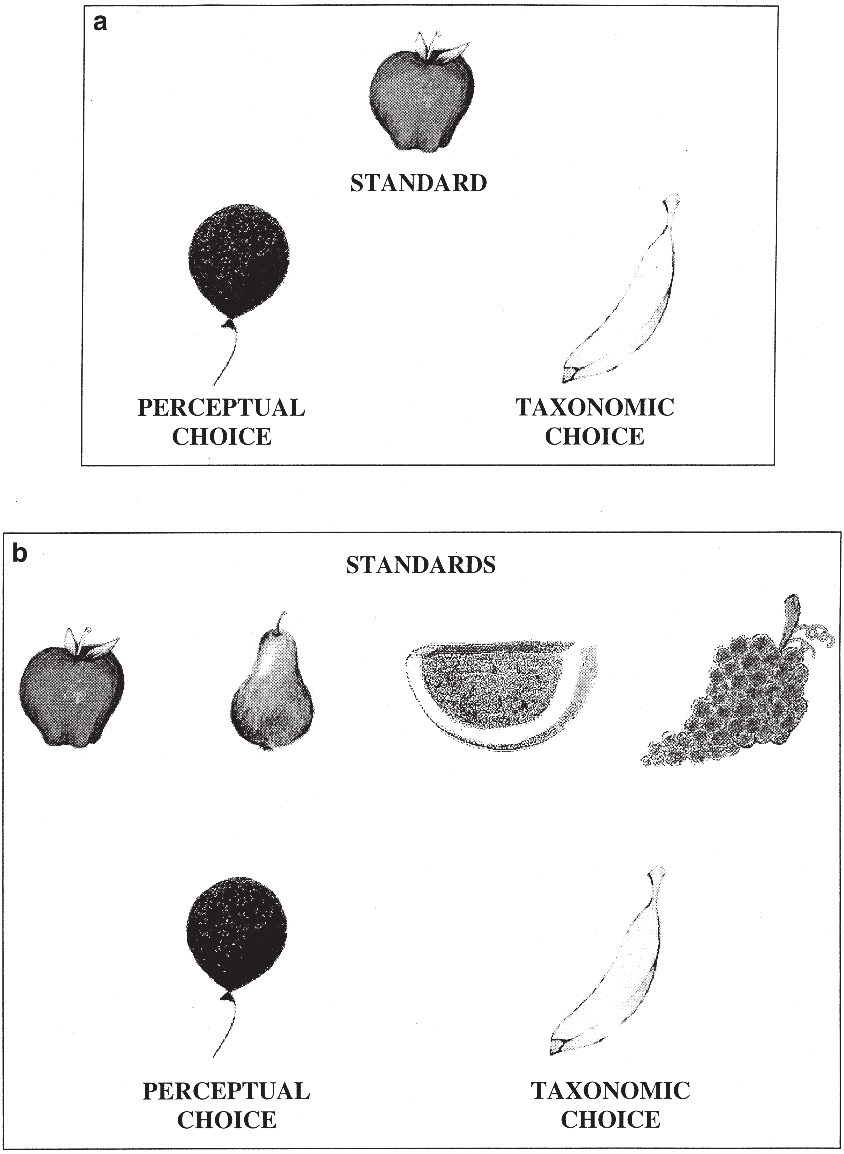

Figure 1. A sample stimulus set from Experiment 1 in the no-compare (1a) and compare (1b) conditions.

In the *compare* condition, the experimenter showed the child *four* standards, placed in a row on an inclined platform, and labeled each of them with the same novel noun: for example, "This is a blicket, and this is a blicket, too, and this one's a blicket, and here's another blicket." For all subjects, the cards were presented in descending order of perceptual similarity to the perceptual alternative as rated by adults. As in the *no-compare* condition, the experimenter asked the child to repeat the novel word and then laid the two alternatives on the table (see Fig. 1b) and asked the child, "Which one of these is a blicket?"

In both conditions, after the child selected one of the choice cards, the experimenter recorded the response and then presented the next trial. If the child did not respond, or if she selected both choice cards, the experimenter pointed to the standard card(s) saying, "Remember, Jo-Jo calls this (these) a blicket. Can you tell Jo-Jo which one of these is a blicket?"

The order of novel words and the item order were varied by selecting two random orders of each, and counterbalancing them within each condition. Left-right placement of the two choice cards was randomly assigned for each trial.

After the child had completed all eight trials, the experimenter again presented each of the stimulus sets and asked the child to tell her the "real" English names for the stimuli, in order to ensure familiarity with the objects depicted. Children's names for the stimuli were evaluated by a rater blind to the experimental condition. All children tested met our inclusion criterion of providing an accurate label, synonym, or functional description for at least 42 of the 48 cards in the stimulus set.

Results

We compared the proportion of trials on which children selected the category choice card in each condition. Children in the *compare* condition ($M = 0.84$, $SD = 0.20$) selected the category choice cards more frequently than those in the *no-compare* condition ($M = 0.39$, $SD = 0.32$), $t(14) = 3.43$, $p < .005$. Performance in the *compare* condition was significantly greater than the chance response rate of .50, $t(7) = 4.95$, $p < .005$. Performance in the *no-compare* condition did not differ from chance responding. Introducing children to multiple members of an object category increases the likelihood that the children will select a conceptual over a perceptual match.

We performed two additional analyses to explore this effect in greater detail. First, we examined individual children's patterns of performance in each condition. Using the binomial formula, with eight trials, an individual child must select category choice cards on at least seven of the eight trials in order to be reliably above chance. We found that five of the eight children in the *compare* condition responded at a rate that exceeded chance, whereas only one of the eight children in the *no-compare* condition was above chance. This difference between the two conditions was marginally reliable according to the Fisher's exact test, $p = .06$. This suggests that the group differences reflect reliably different response patterns across individual subjects within the two conditions.

Our final analysis examined whether the effect of comparison holds up when items, rather than subjects, are used as a random factor. An item analysis also yielded significant differences between the two conditions, $t(7) = 5.81$, $p < .001$. Children's performance on individual sets is listed in Table 2 in the same order as in Table 1. This outcome suggests that the pattern of results is consistent across items as well. In fact, subjects in the *compare* condition selected category choices more frequently than those in the *no-compare* condition for seven of the eight items, which is significant by a sign test, $p < .05$.

Discussion

The results of this study replicate previous findings that children's categorization is affected by surface-level perceptual features in the *no-compare* condition. More importantly for our purposes, this study reveals that experience with multiple instances of an object category leads the children to prefer a taxonomic over a perceptual match. When children see only one standard object, they are equally likely to select either the perceptual or the category match for the standard. However, viewing multiple instances of the target category enabled the children to ignore compelling perceptual commonalities and focus on the deeper, conceptually relevant features of the objects.

These data are consistent with the position that engaging in active comparison across multiple instances of an object category enables children to form conceptual abstractions. However, although this effect argues in a general way for the efficacy of comparisons in category learning, it does not permit specific conclusions as to the nature of the learning process. We would like to conclude that comparison among the exemplars led to a gain in taxonomic insight, but there is a less interesting perception-based model that might account for the data. Children in the *compare* condition received a greater variety of perceptual information within the category than did those in the *no-compare* condition. Some of the exemplars in the *compare* condition may have happened to perceptually favor the category match, whereas the exemplar in the *no-compare* condition clearly favored the perceptual match. Thus the superior performance in the *compare* con-

Table 2. Results of Experiment 1: Children's Proportion of Taxonomic Choices for Each Set, by Condition

	No-compare	Compare
Set 1	0.13	0.88
Set 2	0.38	0.75
Set 3	0.38	1.00
Set 4	0.38	0.88
Set 5	0.50	1.00
Set 6	0.50	0.88
Set 7	0.38	0.88
Set 8	0.50	0.50
Mean	0.39	0.84

dition could have stemmed from simple exemplar-based matching, rather than from active comparison and abstraction.

Another possibility is that simple featural comparison, not structural alignment, is driving the shift from perceptual to category responding. The greater perceptual variety of exemplars in the *compare* condition could have permitted a variety of feature-level matches. A traditional feature-overlap account would thus predict that the *compare* group would show less reliance on any one particular perceptual dimension in transfer. Clearly, if children shifted for these reasons, this would not constitute evidence for a structural alignment process that leads to greater focus on abstract relational structure.

In the next experiment, we set out to differentiate the structural alignment account from a simple perceptual generalization account. We used the same word-extension task as in Experiment 1. The logic of Experiment 2 is to devise two standards, *each* more perceptually similar to the perceptual alternative than to the category alternative. We verify that children will fail to select the category choice when shown *either* of the single standards. Then we test whether a shift towards category responding occurs when both standards are seen together, in the *compare* condition. If such a shift occurs, it cannot be attributed to simple perceptual similarity between the individual items and the alternatives, nor to simple abstraction over common perceptual features in the *compare* condition. For if either of these were driving the effect, then seeing two standards that are both more similar to the perceptual choice than to the category choice should, if anything, *increase* perceptual responding, rather than decrease it.

However, if comparing instances induces a structural alignment process, then children in the *compare* condition may be led to focus on common relational structures that they may typically not notice explicitly—such as how the objects are used and what causal activities they normally participate in. Thus, the prediction of structural alignment theory is that comparison should lead to a shift towards category responding, despite strong perceptual similarity between the standard instances and the perceptual choice.

An additional factor we consider in Experiment 2 is the extent to which the effects of comparison observed in Experiment 1 occur uniquely within word learning contexts. That is, will comparison facilitate category responding even in the absence of a novel word? As noted above, there is evidence that comparison facilitates learning in a variety of contexts, including nonlinguistic contexts (Kotovsky & Gentner, 1996; Loewenstein & Gentner, 1998; Namy et al., 1997). However, prior categorization studies show that children tend to respond more systematically when categorization tasks are framed in terms of word learning. If we find a comparison-based taxonomic shift, then it will be important to assess whether this shift is unique to word-learning contexts. Therefore we included a second factor: whether the children were taught a novel word for the standard(s), or simply given the standards without labels. We predict that comparison will facilitate category responding, even in the absence of a novel word. However, it also seems likely that the effect of comparison may be heightened in a word-learning context.

EXPERIMENT 2

Method

Participants. Eighty 4-year-olds (mean age = 4;4, range = 3;6–4;11 participated. Thirteen additional children were excluded, seven for failure to accurately label enough stimulus items during a post-experiment naming task (see description of inclusion criterion below), two due to experimenter error, and four due to failure to complete the task. The children were drawn from the same population as was used in Experiment 1. Half the children participated in the *word* condition and half in the *no-word* condition.[2]

Materials. Forty colored line drawings of real objects were used as stimuli. These drawings were organized into ten sets of four cards each. Each stimulus set included two standards and two choice alternatives. The two standards were perceptually similar members of an object category (e.g., a bicycle and a tricycle). As in Experiment 1, the two alternatives included a perceptual alternative—a pictured object that was perceptually similar to the standards but conceptually unrelated (e.g., a pair of glasses)—and a category alternative—an object from the same category as the standards, but perceptually quite distinct from them (e.g., a skateboard). The ten item sets included items from the eight sets used in Experiment 1, plus two additional sets. Table 3 gives the complete list of stimuli and Figure 2 shows an example stimulus set.

Table 3. Materials Used in Experiment 2

			Alternatives	
	Standard 1	**Standard 2**	**Perceptual**	**Taxonomic**
1.	Apple	Pear	Balloon	Banana
2.	Plate	Bowl	Cookie	Casserole Dish
3.	Drum	Tambourine	Hat Box	Flute
4.	Carrot	Corn	Rocket	Turnip
5.	Ice Cream	Lollipop	Top	Candy Bar
6.	Baseball Cap	Derby Hat	Igloo	Sombrero
7.	Bicycle	Tricycle	Glasses	Skateboard
8.	Caterpillar	Snake	Rope	Turtle
9.	Baseball Bat	Golf Club	Pencil	Tennis Racket
10.	Baseball	Beach Ball	Orange	Football

[2] Assignment was not entirely random, in that 24 of the 40 children in the *word* condition were run as a separate sample prior to the collection of data from the *no-word* and the remaining 16 *word* children. However, the children in both samples were recruited from the same cohort of families, and performance in the *word* condition did not differ as a function of sample. We therefore report the data in a pooled format.

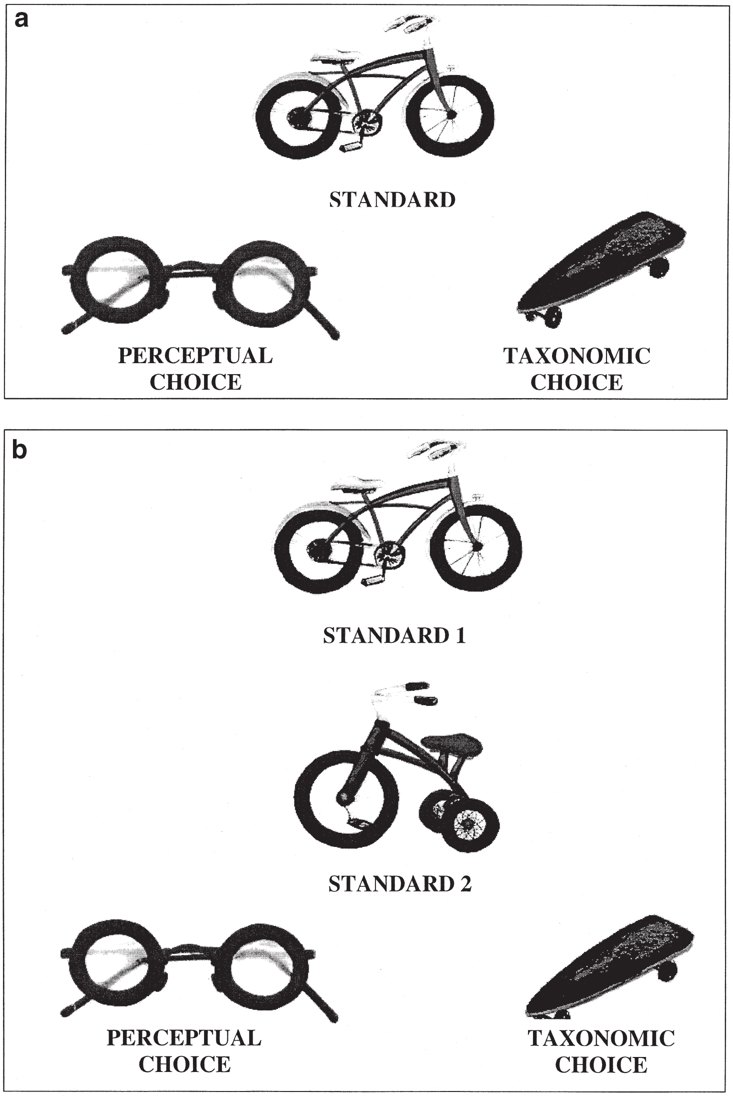

Figure 2. **A sample stimulus set from Experiment 2 in the no-compare (2a) and compare (2b) conditions.**

Stimulus selection. The two standards used in this experiment were the two standards from Experiment 1 that had been rated by adults as highest in similarity to the perceptual choice in the first experiment. These standards were highly similar to each other and were consistently rated as more similar to the perceptual choice than to the category choice. Adults also rated the similarity of the two additional sets, and the same relationship among the stimuli existed for these sets.

Procedure. The procedure was similar to that of Experiment 1. Participants were seated at a low table in the laboratory across from the experimenter. Parents were seated behind the children and were explicitly instructed not to speak to the children during testing. As in Experiment 1, children were introduced to Jo-Jo, a toy stuffed dog. Those learning novel words were told that they would learn Jo-Jo's special names for things. Those in the *no-word* control group were told they would be playing a special game with Jo-Jo.

The children were assigned to either the *compare* or the *no-compare* condition, and either the *word* or *no-word* group. Among children learning novel words, the procedure in the *no-compare* condition was identical to Experiment 1, with the exception that half the children received one of the two standards used in each set, and the other half received the other (see Fig. 2a). In the *compare* condition, the experimenter showed the child *two* standard cards, one placed directly above the other, and labeled each with a novel noun, for example, "This is a blicket, and this is a blicket, too." The order in which the two cards were presented was counterbalanced across subjects. The experimenter asked the child to repeat the novel word. She then laid the two choice cards on the table, below and to either side of the standards (see Fig. 2b) and asked the child, "Which one of these is a blicket?"

In both conditions, after the child selected one of the choice cards, the experimenter recorded the response and then presented the next trial. If the child did not respond, or if she selected both choice cards, the experimenter pointed to the standard card(s) saying, "Remember, Jo-Jo calls this a blicket. Can you tell Jo-Jo which one of these is a blicket?"

For children in the *no-word* group, the procedure was identical, but the wording used by the experimenter differed. In the *no-compare* condition, the experimenter directed children's attention to the target without labeling it, saying, "Look at this one!" She then placed the two choice cards on the table asking, "Can you find another one that's the same kind as this?" In the *compare* condition, the experimenter presented the two standards and highlighted their commonalities without naming the objects, saying, "See this one, and see this one? *See how these are the same kind of thing*?" She then placed the choice cards on the table, saying, "Can you find another one that's the same kind as these?"

The order of novel words and the item order were varied by selecting two random orders of each, and counterbalancing them within each condition. Left-right placement of the two choice cards was randomly assigned for each trial.

After the child had completed all 10 trials, the experimenter again presented each of the stimulus sets and asked the child to tell her the "real" English names for the stimuli, in order to ensure familiarity with the objects depicted. Children's names for the stimuli were evaluated by a rater blind to the experimental condition. Children who did not provide an accurate label, close synonym, or functional description for at least 35 of the 40 cards in the stimulus set were excluded from the analysis.

Results and Discussion

As predicted, children in the *compare* condition selected category choices more often than those in the *no-compare* condition. Figure 3 shows the mean proportions of category responding in each group and Table 4 shows the results by item. A 2 (Condition: *compare* vs. *no-compare*) × 2 (Word: *word* vs. No-word) ANOVA confirmed a main effect of condition, $F_s(1, 76) = 4.59$, $p < .05$, $F_i(1, 9) = 28.43$, $p < .001$. Thus, introducing children to two perceptually similar members of an object category increases the likelihood that the children will select a conceptual over a perceptual match, even when *both* standards are individually more similar to the perceptual match than to the conceptual match. These findings run strongly counter to the standard similarity-generalization account.

Turning to the question of labeling effects, the analysis revealed a main effect of word in the item ANOVA, $F_i(1, 9) = 5.43$, $p < .05$, although not in the subject ANOVA, $F_s(1, 76) = .96$, *ns.* Contrary to expectation, there was no interaction between comparison condition and word. Thus we found no evidence that the effects of comparison are specific to word-learning contexts.

Despite the non-significance of the Condition X Word interaction, the trend of the data suggests some influence of language on comparisons. Therefore, we ex-

Table 4. Results of Experiment 2: Children's Proportion of Taxonomic Choices for Each Set, by Condition

	Word		No-word	
	No-compare	**Compare**	**No-compare**	**Compare**
Set 1	0.50	0.60	0.35	0.50
Set 2	0.30	0.75	0.35	0.35
Set 3	0.50	0.65	0.25	0.40
Set 4	0.35	0.65	0.30	0.50
Set 5	0.45	0.55	0.25	0.45
Set 6	0.30	0.75	0.70	0.55
Set 7	0.65	0.75	0.55	0.65
Set 8	0.65	0.65	0.50	0.65
Set 9	0.35	0.35	0.40	0.35
Set 10	0.45	0.45	0.40	0.60
Mean	0.46	0.62	0.41	0.50

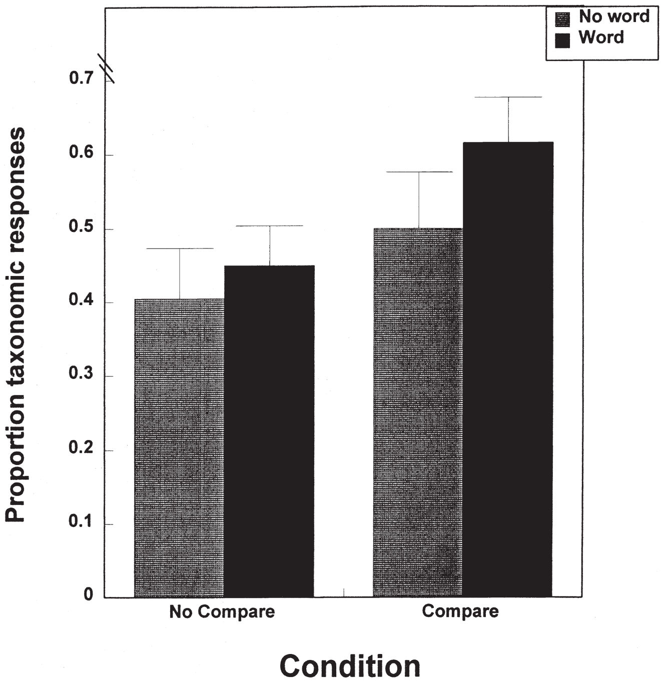

Figure 3. A results of Experiment 2: mean proportion of taxonomic choices as a function of condition and word assignment.

amined individual children's patterns of performance in each condition. Using the binomial formula, with ten trials, children must select category choice cards on at least eight of the ten trials in order to be reliably above chance. We found that among children learning words, 7 of the 20 children in the *compare* condition responded at above-chance rates, as compared to only 2 of the 20 children in the *no-compare* condition. The number of children who selected the category choice card at above chance rates differed marginally as a function of condition, according to the Fisher's exact test, $p = .054$. In contrast, in the *no-word* group, 5 of the 20 children in the *compare* condition and 3 of the 20 children in the *no-compare* condition responded at above-chance rates. This difference was not reliable, $p = .23$. These results are not conclusive, but they provide suggestive evidence that receiving the same label for a pair of items helps invite the comparison process.

Parts analysis. Finally, we considered an alternative perceptual-based account for these effects. Perhaps the two standards shared particular parts or features that were also shared uniquely with the taxonomic but not the perceptual choice. For example, in Figure 2 the bicycle and the tricycle shared wheels with the skateboard (the category alternative) but not with the glasses (the perceptual alternative). In this case, children might have benefited from comparison at a more shallow perceptual level; placing two objects with a shared part in close proximity may have heightened the salience of that particular part. This might have led children to categorize on the basis of this part.

To ascertain whether this explanation might account for the heightened category responding in the *compare* condition, we (1) performed a parts analysis of our stimuli, and (2) examined whether particular common parts could have influenced performance on the forced-choice task. We asked two adult raters, naïve to the experimental hypotheses, to inspect the two standards in each set (in the absence of the choice cards) and list any shared features between them. After they had done so, we presented the two choice cards and asked the raters to evaluate (using a 1 to 5 rating scale) the extent to which the listed features were shared by either of the choice cards.

Using a liberal criterion, we counted a set as presenting a potential part confound if the category choice was given a mean rating of 4 or higher *and* the perceptual choice was given a rating of 3 or lower. This criterion indicated that several sets, including sets 1, 3, 5, 8, 9, and 10 may have included a potential confound in that a particular feature was shared by the two standards and the category choice but not the perceptual choice. We examined whether children were more likely to select category choices on these items that possessed a featural confound. Interestingly, children in the *compare* condition were actually *less* likely to select the taxonomic choice on the sets in which the taxonomic choice shared parts with the standards ($M = .54$, $SD = .29$) than on those sets without shared features ($M = .73$, $SD = .29$), $t(19) = 4.26$, $p < .001$. Thus, the matching of individual parts cannot account for the effects of comparison on taxonomic responding.

In this experiment, as in the first experiment, children in the *compare* condition were likely to select the perceptually dissimilar same-category items as a match for the standard category, whereas those in the *no-compare* condition were equally likely to choose the taxonomic or perceptual match. Thus, when children are introduced to two highly similar members of a category, they are *less* likely to select a perceptual match and *more* likely to select a same-category match, even if it is perceptually disparate from the standards. This effect held regardless of whether the category was labeled (although language appeared to contribute to the effect, as discussed below). This outcome is counterintuitive from the point of view of the shallow similarity-generalization account of category learning. However, this result is precisely what the structural alignment account of categorization would predict. Co-categorizing and comparing two alignable objects can lead children to notice and use deeper structural commonalities that are not readily evident from the perceptual information.

General Discussion

In the two studies presented here, we tested whether comparisons across items helped children to arrive at conceptual understanding of categories. Specifically, we hypothesized that the process of structural alignment would promote the discovery of relatively abstract relational commonalities that could characterize the category being learned. In both studies, children taught a new word for an object were as likely to extend the word to a perceptually similar object as to a conceptually similar object. We asked whether comparison to other exemplars would lead children to consistently choose the conceptually similar extension. In Experiment 1, we found that category-based extensions were more likely when children saw multiple exemplars of the new word that varied in their perceptual characteristics. These results are consistent with the predictions of structural alignment. However, they could equally well be explained in terms of perceptual feature overlap between individual standard items and the two response alternatives, or with simple perceptual generalization across the set of standards.

In Experiment 2, we sought more specific evidence for structural alignment processing in the contribution of comparison to category learning. We compared learning one example of the category with learning two examples; moreover we designed the comparison pairs so that both members were more similar perceptually to the high-similarity out-of category alternative than to the same-category alternative. As in Experiment 1, children were (nonsignificantly) more likely to chose the perceptual alternative when extending a name for either of the initial exemplars alone. But when the two exemplars were juxtaposed and given the same name, children were far more likely to choose the category item as a new exemplar of the word. This cannot be a simple "breadth of training" or feature overlap effect, because both members of the training set by themselves tend to be extended to the perceptual alternative (and are rated by adults as more perceptually similar to it). These findings are strong evidence for an alignment process by which children who may initially notice and align on the basis of perceptual similarity are led to apprehend other commonalities, less obvious but often deeper and more conceptually connected.

We began this research by posing the question of how to reconcile two seemingly contradictory generalizations. The first is that language—word meaning in particular—provides a major avenue through which children gain understanding of conceptual structure. The second is that children's early word meanings, as evidenced by their extensions to new exemplars, are heavily influenced by perceptual commonalities such as common shape. We have proposed that the process of structural alignment may act as a bridge from an initial perceptually-based category to a later more sophisticated understanding of the category.

The Role of Similarity in Conceptual Development

In the behaviorist tradition, similarity—especially perceptual similarity—was considered a major force in learning. Cognitive theorists have noted that learning via perceptual similarity cannot account for the level of abstract knowledge achieved by human learners (Gelman & Coley, 1990; Keil, 1989; Murphy & Medin, 1985; Rips, 1989). As Quine (1969) puts it, there is little reason to think that "the muddy old notion of similarity" (p. 172) has anything to contribute to the development of abstract capacities. Indeed, in the conceptual development literature, perceptual similarity has sometimes been viewed as a deceiver, a tempting impostor that leads children to settle for merely superficial resemblances instead of seeking out theory-based explanations for category structure. Such a view, we suggest, undervalues the informative potential of the comparison *process*. The implicit assumption is that similarity processing is self-terminating, stopping as soon as a clear likeness is identified. But although this may possibly be true in some contexts—perhaps especially in certain concept-learning paradigms—in ordinary life, noticing one likeness invites noticing others.

Research in analogy and similarity suggests that the comparison process can change knowledge in at least three ways: by highlighting commonalties, by inviting inferences from the more familiar to the less familiar item, and by suggesting adaptations or re-representations of the two concepts that increase the degree of alignment (Gentner & Markman, 1997; Gentner & Medina, 1998; Gentner & Wolff, in press; see also Gick & Holyoak, 1983; Goldstone, 1994; Holyoak, Novick, & Melz, 1994; Karmiloff-Smith, 1991). Juxtaposition of two perceptually similar exemplars, especially in the presence of a common label, may invite the child to notice further, more abstract, commonalities (Gentner & Rattermann, 1991; Kotovsky & Gentner, 1996). These more abstract properties, once noticed, could be used to determine other category members without needing the support of perceptual similarity. We suggest that in word learning, comparisons between and among exemplars and the word's current representation may continue to enrich the child's representation of the concept long after the initial meaning is formed.

Additional evidence that perceptual similarity may act to bootstrap children's understanding of deeper commonalities comes from Kemler Nelson's (1995) study. Recall that Landau et al. (1998) found that 2-year-olds extended word meanings on the basis of shape, even when a cross-cutting salient function was clearly demonstrated. However, Kemler Nelson (1995) found that 3- to 6-year-old children extended word meanings on the basis of function, as well as perceptual similarity. That Kemler Nelson's 3-year-olds were better able to go beyond perceptual likeness than Landau et al.'s 2-year-olds may, of course, be simply a function of level of learning and development. However, there is another possible contributing factor. This function-oriented pattern occurred only in cases where the part of the object that was critical for its function was perceptually salient relative to its overall appearance. When this was not the case, children extended the word solely on the basis of perceptual similarity. Kemler Nelson noted that a convergence between common perceptual structure and common functional affordance appeared to facilitate children's focusing on the functional commonalities.

The Role of Language

Prior evidence suggests that common labels can invite alignment. Waxman and Markow (1995) have found evidence for such a phenomenon in young infants. At 12 months of age, infants habituated to a series of highly perceptually similar instances from the same basic level category (such as cats), but did not habituate to a less similar series of instances from a broader set of mammals. However, when the same novel noun was applied to each instance, the infants habituated even to this less similar series. There is other research suggesting that alignment at the basic level, for which perceptual similarity is extremely high, may be a particularly effortless process (e.g., Markman & Wisniewski, 1997). Many of children's first words are basic-level nouns (Gentner & Boroditsky, in press; Rosch et al., 1976; however, see Nelson, Hampson, & Kessler-Shaw, 1993). Quinn and Eimas (1996) have shown habituation to basic level categories in infants as young as 3 months of age. Golinkoff et al. (1995) found that in a word-extension task, 3- and 4-year-olds could extend novel nouns at the basic level, but not at the superordinate level, suggesting that early word understanding may rely on perceptual as well as conceptual commonalities.

It is often argued that words serve as invitations to seek out and form meaningful conceptual categories (Balaban & Waxman, 1997; Brown, 1958; Gelman & Coley 1990; Gelman & Wellman, 1991; Waxman & Markow, 1995). As Gentner and Rattermann (1991, p. 260) put it, "a word can function as a promissory note, signaling subtle commonalities that the child does not yet perceive." Our results were partly supportive of this claim. It appears that the use of common labels increased category responding. However, our expectation that common language would heighten the effect of comparison was only weakly supported. There was no significant interaction between comparison condition and word–no-word condition, although the chance analyses suggest a borderline trend towards stronger comparison effects in the common label condition. Namy and Gentner (1999) find additional evidence to support this trend. We find that providing a common label for two standards facilitates attention to category relations, but that providing two contrasting labels for the same two standards inhibits category abstraction. This outcome supports the claim that patterns of linguistic labeling guide children's comparison processes.

Issues for Future Research

In our studies, we found insightful category generalization when the two standards: (1) belonged to the same category, (2) were given a common label, and (3) had high perceptual similarity. A natural question for further research is what would happen if we altered one or another of these conditions. Namy and Gentner (1999) have found that if condition (1) is suspended—that is, if a common label is applied to two standards that do not belong to the same category, so that no deeper alignment is possible—then children continue to show perceptual responding. Namy and Gentner also tested the importance of condition (2), a common label, as noted above. Finally, we are currently investigating condition (3) by varying the perceptual similarity relation among the standards.

Summary

To return to our initial question, a resolution of the paradox of early word learning might take the following form. Strong perceptual similarity, which is easily detected even by “universal novices” such as very young children, creates an invitation to alignment. The resulting comparison process can promote the noticing of further commonalties. If the match is conceptually justified—that is, if the “kind world” assumption holds, as it does for basic-level concepts—then the comparison process is likely to reveal further functional and relational commonalities. These often work in tandem with perceptual likeness, but they may also come to outweigh the perceptual commonalities. Categories may then be assigned on grounds of causal or functional likeness rather than by perceptual similarity, as when children come to define “island” as a body of land surrounded by water, rather than as a warm and sandy place (Keil, 1989; see also Carey, 1985; Gentner & Rattermann, 1991; Rips, 1989, 1991).

If these conjectures are correct, then even a mechanism as simple as the shape bias can play a useful role in development. It allows young children to discover and entertain a set of candidate extensions, many of which will turn out to be reasonable, especially in cases where perceptual properties are correlated with relation properties. Indeed, the prevalence of basic-level terms in early vocabularies may be a result of this natural pattern. Paradoxically, the very comparison process initiated by strong perceptual similarity may facilitate the eventual ability to go beyond perceptual to conceptual likeness in word extension. In our studies, a comparison process among highly similar objects encouraged the noticing of deeper commonalities among objects in preschoolers’ categorization. These findings demonstrate how the process of comparison, viewed as a process of structural alignment, constitutes an important route towards abstract conceptual understanding.

References

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

Baldwin, D. A. (1989) Priorities in children’s expectations about object label reference: Form over color. *Child Development, 60*, 1291–1306.

Baldwin, D. A. (1992). Clarifying the role of shape in children’s taxonomic assumption. *Journal of Experimental Child Psychology, 54*, 392–416.

Brown, R. (1958). *Words and things: An introduction to language*. New York: Free Press.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

Clark, E. (1973) What’s in a word? On the child’s acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 65–110). New York: Academic Press.

Deloache, J. S. (1995). Early understanding and use of symbols—the model model. *Current Directions in Psychological Science, 4*, 109–113.

Elio, R., & Anderson, J. R. (1984). The effects of information order and learning mode on schema abstraction. *Memory and Cognition, 12*, 20–30.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41,* 1–63.

Forbus, K., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science, 19*(2), 141–205. (Abridged version to be reprinted in *Cognitive Modeling,* by T. Polk & C. M. Seifert, Eds., in press, Boston: MIT Press)

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology, 26,* 796–804.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23,* 183–209.

Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the nonobvious. *Cognition, 38,* 213–244.

Gentner, D. (1978). What looks like a jiggy but acts like a zimbo? A study of early word meaning using artificial objects. *Papers and Reports on Child Language Development, 15,* 1–6.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7,* 155–170. (Reprinted in *Readings in cognitive science: A perspective from psychology and artificial intelligence (1988)* by A. Collins & E. E. Smith (Eds.), Palo Alto, CA: Kaufmann)

Gentner, D., & Boroditsky, L. (in press). Individuation, relativity, and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development.* New York: Cambridge University Press.

Gentner, D., & Imai, M. (1995). A further examination of the shape bias in early word learning. *Proceedings of the Twenty-sixth Annual Child Language Research Forum* (167–176). Stanford, CA.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52,* 45–56.

Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition, 65,* 263–297.

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. Gelman (Ed.), *Language and conceptual development* (pp. 225–277). New York: Cambridge University Press.

Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability and inferential soundness. *Cognitive Psychology, 25,* 524–575.

Gentner, D., & Wolff, P. (in press). Metaphor and knowledge change. In E. Dietrich & A. B. Markman (Eds.), *Cognitive dynamics: Conceptual and representational change in humans and machines.* Mahwah, NJ: Erlbaum.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15,* 1–38.

Goldstone, R. L. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition, 52,* 125–157.

Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology, 23,* 222–264.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology, 28,* 99–108.

Golinkoff, R. M., Shuff-Bailey, M., Olguin, R., & Ruan, W. (1995). Young children extend novel words at the basic level: Evidence for the principle of categorical scope. *Developmental Psychology, 31*, 494–507.

Holyoak, K. J., Novick, L. R., & Melz, E. R. (1994). Component processes in analogical transfer: Mapping, pattern completion, and adaption. In K. J. Holyoak & J. A. Barnden (Eds.), *Advances in connectionist and neural computation theory: Vol. 2. Analogical connections* (pp. 113–180). Norwood, NJ: Ablex.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science*, *13*, 295–355.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427–466.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development, 9*, 45–75.

Kalish, C. W., & Gelman, S. A. (1992). On wooden pillows: Multiple classifications and children's category-based inductions. *Child Development, 63*, 1536–1557.

Karmiloff-Smith, A. (1991). Beyond modularity: Innate constraints and developmental change. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 3–36). Hillsdale, NJ: Erlbaum.

Keane, M. T., & Brayshaw, M. (1988). The incremental analogical machine: A computational model of analogy. In D. Sleeman (Ed.), *Third European Working Session on Machine Learning* (pp. 53–62). San Mateo, CA: Kaufmann.

Keil, F. (1989). *Concepts, kinds, and conceptual development.* Cambridge, MA: MIT Press.

Kemler Nelson, D. G. (1995). Principle-based inferences in young children's categorization: Revisiting the impact of function on the naming of artifacts. *Cognitive Development, 10*, 347–380.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*, 2797–2822.

Landau, B., Smith, L. B., & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*, 299–321.

Landau, B., Smith, L. B., & Jones, S. (1998). Object shape, object function, and object name. *Journal of Memory & Language, 38*, 1–27.

Loewenstein, J. (1997, April). *Using comparison to improve preschooler's spatial mapping ability.* Paper presented at the Society for Research in Child Development, Washington, DC.

Loewenstein, J., & Gentner, D. (1998). Relational language facilitates analogy in children. *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*(pp. 615–620). Mahweh, NJ: Lawrence Erlbaum Associates.

Loewenstein, J., & Gentner, D. (1999). *Structural alignment promotes the development of mapping.* Manuscript under review.

Madole, K. L, Oakes, L. M., & Cohen, L. B. (1993). Developmental changes in infants' attention to function and form-function correlations. *Cognitive Development, 8*, 189–209.

Mandler, J. M. (1992). How to build a baby II: Conceptual primitives. *Psychological Review, 99*, 587–604.

Mandler, J. M., & Bauer, P. J. (1988). The cradle of categorization: Is the basic level basic? *Cognitive Development, 3*, 247–264.

Mandler, J. M., Bauer, P. J., & McDonaugh, L. (1991). Separating the sheep from the goats: Differentiating global categories. *Cognitive Psychology, 23*, 263–298.

Markman, A. B., & Gentner, D. (1993a). Splitting the differences: A structural alignment view of similarity. *Journal of Memory & Language*, *32*, 517–535.

Markman, A. B., & Gentner, D. (1993b). Structural alignment during similarity comparisons. *Cognitive Psychology, 25,* 431–467.

Markman, A. B., & Gentner, D. (1996). Commonalities and differences in similarity comparisons. *Memory & Cognition, 24,* 235–249.

Markman, A. B., & Wisniewski, E. (1997). Similar and different: The differentiation of basic-level categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 54–70.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction.* Cambridge, MA: MIT Press.

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

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100,* 254–278.

Medin, D. L., & Ross, B. H. (1989). The specific character of abstract thought: Categorization, problem solving and induction. *Advances in the Psychology of Human Intelligence, 5,* 189–223.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92*, 289–316.

Namy, L. L., & Gentner, D. (1999). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. Manuscript in preparation.

Namy, L. L., Smith, L. B., & Gershkoff-Stowe, L. (1997). Young children's discovery of spatial classification. *Cognitive Development, 12*, 163–184.

Nelson, K. (1973). Some evidence for the cognitive primacy of categorization and its functional basis. *Merrill-Palmer Quarterly, 19*, 21–39.

Nelson, K., Hampson, J., & Kessler-Shaw, L. (1993). Nouns in early lexicons: Evidence, explanations, and implications. *Journal of Child Language, 20*, 61–84.

Oakes, L. M., Plumert, J. M., Lansink, J. M., & Merryman, J. D. (1996). Evidence for task dependent categorization in infancy. *Infant Behavior and Development, 19*, 425–440.

Quine, W. V. (1969). *Ontological relativity and other essays.* New York: Columbia University Press.

Quinn, P. C., & Eimas, P. D. (1996). Perceptual cues that permit categorical differentiation of animal species by infants. *Journal of Experimental Child Psychology, 63,* 189–211.

Ricciuti, H. N. (1965). Object grouping and selective ordering in infants 12 to 24 months old. *Merill-Palmer Quarterly, 11,* 129–148.

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.

Rips, L. J. (1991). Similarity and the structure of categories. In D. J. Napoli & J. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore festschrift for Lila Gleitman.* Hillsdale, NJ: Erlbaum.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Ross, G., Nelson, K., Wetstone, H., & Tanouye, E. (1986). Acquisition and generalization of novel object concepts by young language learners. *Journal of Child Language, 13*, 67–83.

Skorstad, J., Gentner, D., & Medin, D. L. (1988). Abstraction processes during concept learning: A structural view. *Proceedings of the Tenth Annual Conference of the Cognitive Science Society* (pp.419–425). Hillsdale, NJ: Erlbaum.

Smith, L. B., Jones, S. S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretations. *Developmental Psychology, 28*, 273–286.

Sugarman, S. (1983). *Children's early thought: Developments in classification.* New York: Cambridge University Press.

Vygotsky, L. S. (1962). *Thought and language.* Cambridge, MA: MIT Press.

Ward, T. B., Vela, E., Peery, M. L., Lewis, S. N., Bauer, N. K., & Klint, K. A. (1989). What makes a vibble a vibble? A developmental study of category generalization. *Child Development, 60*, 214–224.

Waxman, S. R. (1990). Linguistic biases in the establishment of conceptual hierarchies: Evidence from preschool children. *Cognitive Development, 5*, 123–150.

Waxman, S. R., & Gelman, R. (1986). Preschoolers' use of superordinate level relations in classification and language. *Cognitive Development, 1*, 139–156.

Waxman, S. R., & Hall, D. G. (1993). The development of a linkage between count nouns and object categories: Evidence from 15- to 21-month-old infants. *Child Development, 64*, 1224–1241.

Waxman, S. R., & Kosowski, T. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. *Child Development, 61*, 1461–1473.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*, 257–302.