Infant and Child Development
Inf. Child Dev. **22**: 1–23 (2013)
Published online 2 July 2012 in Wiley Online Library
(wileyonlinelibrary.com). DOI: 10.1002/icd.1764

Comparison Versus Contrast: Task Specifics Affect Category Acquisition

Amber A. Ankowski*, Haley A. Vlach and Catherine M. Sandhofer
Department of Psychology, University of California, Los Angeles, CA, USA

A large literature has documented that comparison and contrast lead to better performance in a variety of tasks. However, studies of comparison and contrast present contradictory conclusions as to when and how these processes benefit learners. Across four studies, we examined how the specifics of the comparison and contrast task affect performance by systematically manipulating the feature variation and category structure in a category extension task performed by 3-year-old and 4-year-old children. Studies 1 ($n = 48$, $M = 42.6$ months) and 2 ($n = 48$, $M = 42.$ 4 months) investigated comparison and contrast with high-density categories. Studies 3A ($n = 60$, $M = 43.47$ months), 3B ($n = 48$, $M = 53.$ 2 months) and 4 ($n = 48$, $M = 53.7$ months) investigated comparison and contrast with low-density categories. Results indicated both category structure and feature variation affect the efficacy of comparison and contrast. Copyright © 2012 John Wiley & Sons, Ltd.

Key words: comparison; categorization; cognitive development

Extensive research indicates that providing learners with opportunities to compare (view multiple examples from the target category) or contrast (view a member of the target category with non-members of the category) facilitates categorization and aids the discovery of deeper relational structure (Anggoro, Gentner, & Klibanoff, 2005; Boroditsky, 2007; Gentner, Loewenstein, & Hung, 2007; Gentner, Loewenstein, & Thompson, 2003; Gentner & Namy, 1999; Gick & Holyoak, 1980, 1983; Kurtz, Miao, & Gentner, 2001; Loewenstein & Gentner, 2001; Loewenstein, Thompson, & Gentner, 1999; Namy & Gentner, 2002; Namy, Smith, & Gershkoff-Stowe, 1997; Oakes & Ribar, 2005; Rittle-Johnson & Star, 2007). As a whole, these studies overwhelmingly indicate that experiencing instances simultaneously rather than sequentially strongly benefits learning. However, at the same time, studies of comparison and contrast present a puzzling contradiction: in some studies, contrast appears more effective than comparison in promoting categorization (e.g. Andrews, Livingston, & Kurtz, 2005; Hampton, Estes, & Simmons, 2005; Kalish & Lawson, 2007; Kurtz & Boukrina, 2004), whereas in other studies, comparison appears to more strongly promote

*Correspondence to: Amber A. Ankowski, Department of Psychology, University of California, Los Angeles. E-mail: aankowski@ucla.edu

categorization (e.g. Kotovsky & Gentner, 1996; Namy & Gentner, 2002). In the current series of studies, we examine this contradiction and propose that the efficacy of the learning processes is mediated by specifics of the task. By providing a systematic examination of how task specifics affect comparison and contrast processes, we seek to inform how the processes themselves operate.

COMPARISON AND CONTRAST

Comparison and contrast provide different information about categories. Comparing multiple members of the same category provides information about 'category membership'. For example, comparing multiple members of the category 'red' (e.g. a red ball, a red fire truck and a red apple) provides information about the shared and unshared features between category members. However, contrasting a category member against a non-category member (e.g. contrasting a red object with a yellow object) provides information about 'category boundaries'. That is, contrast provides information about what distinguishes category members from non-members.

Studies that have included both comparison and contrast manipulations have reported mixed results regarding the relative efficacy of the processes. Some studies suggest that comparison more strongly promotes categorization, particularly when the target category is not bound by salient perceptual commonalities (Kotovsky & Gentner, 1996; Namy & Gentner, 2002). For example, Kotovsky and Gentner (1996) showed that young children could recognize higher order commonalities (e.g. symmetry) if they were initially presented with category examples that shared perceptual similarity. Gentner and colleagues have argued that comparing similar category members is critical for category discovery because aligning shared perceptual features between representations allows children to discover underlying higher order commonality (Gentner & Markman, 1994; Gentner & Namy, 1999). Further, the effect of comparison is enhanced if the examples are highly similar. High similarity between exemplars appears to facilitate the detection of 'both' similarities and differences (Gentner & Markman, 1994), thus category acquisition may be better facilitated through comparison of similar category examples than through contrast of dissimilar examples.

In other studies, contrast appears more effective than comparison in promoting categorization (Andrews et al., 2005; Hampton et al., 2005; Kalish & Lawson, 2007; Kurtz & Boukrina, 2004). For example, in one study, adult participants were asked to learn the labels for three types of novel sea creatures (Andrews et al., 2005). Adults were presented with the sea creatures in one of the three between-subjects conditions: (i) a 'compare' condition in which adults always saw category members presented together (e.g. comparing three 'zofs'); (ii) a 'contrast' condition in which adults always saw category members presented with non-category members (e.g. contrasting one 'zof' with one 'gik' and one 'wug'); or (iii) a 'compare–contrast' condition in which adults always both compared category members and contrasted category non-members (e.g. comparing two 'zofs' and contrasting them with one 'gik'). Results showed that the participants were more successful in the contrast condition than in any other condition. Thus, learning about category boundaries through contrasting exemplars proved most important for adult categorization. Further, evidence shows that contrast may especially promote categorization for young children. Kalish and Lawson (2007), for example, found that preschool-aged children relied more on contrastive information than school-aged children and adults in a category extension task.

DOI: 10.1002/icd

We propose that contradictory results as to the effectiveness of comparison and contrast processes may arise because the processes are strongly mediated by specific aspects of the categorization task. Studies of adults' and children's categorization have established that aspects of the task context may have substantial effects on category learning (e.g. Kovack-Lesh & Oakes, 2007; Yeh & Barsalou, 2006). Thus, we suggest that it is not that one process is inherently more effective than the other; instead, task context may moderate children's ability to use comparison and contrast. The present studies investigate how two aspects of task context, the structure of the category as a whole and the amount of irrelevant variation in the category examples, affect the efficacy of comparison and contrast.

The stimuli, and the specific features of those stimuli, have varied widely in studies of comparison and contrast. Stimuli have included model rooms (Loewenstein & Gentner, 2001), fictitious sea creatures (Andrews et al., 2005), familiar objects and animals (Oakes & Ribar, 2005; Waxman & Klibanoff, 2000) and paint samples (Hampton et al., 2005). Additionally, the stimuli have varied in the degree to which they include features that are irrelevant to the categorization task. For example, Oakes and Ribar (2005) presented infants with different pairs of cats that varied in features that did not affect membership to the category of cats (e.g. fur length and fur color). This type of variation is typically unstudied (see Kovack-Lesh & Oakes, 2007 for a discussion); however, these irrelevant features may indeed be important to categorization because the overall similarity of the exemplars may contribute to children's ability to extract relevant category information (Gentner & Markman, 1994). Thus, the amount of commonality between exemplars, both in features that define category membership and those that do not, may influence category acquisition.

A second reason that comparison and contrast research may yield contradictory results is that these processes may work better with some category structures than others. Previous studies have presented participants with a diverse range of category structures, including categories in which the category members shared a high degree of overall similarity (e.g. cats: Oakes & Ribar, 2005) and categories in which the category members shared a smaller degree of overall similarity (e.g. taxonomic categories: Namy & Gentner, 2002; spatial relations: Loewenstein & Gentner, 2001). Highlighting similarity between category members may benefit categories in which the members share relatively little perceptual similarity, whereas highlighting similarity between category members may be superfluous for categories in which the members already share a high degree of perceptual similarity.

In the current studies, we manipulated category structure by presenting children with either high-density or low-density categories (Kloos & Sloutsky, 2008). For both category structures, the categories to-be-learned were defined by the lightness or darkness of color (i.e. dark, medium, light). 'High-density' categories have category members that share many common features relevant for category membership and have less variation in features that are irrelevant for category membership. 'Low-density' categories have category members that share fewer common features relevant for category membership and have more variation in features that are irrelevant for category membership. Although both high-density and low-density categories may vary in some dimensions irrelevant for category membership, their key difference lies in the number of dimensions predictive of category membership. Kloos and Sloutsky (2008) found that the statistical density of categories affected the ease of acquisition, such that more features jointly predictive of category membership made categorization easier.

DOI: 10.1002/icd

Higher density categories were more readily acquired, whereas lower density categories posed a greater challenge to learners. Because acquiring low-density categories can be difficult, even for adults, they may be a particular challenge for young learners.

The high-density categories in the present study contained category members that shared perceptual commonality in multiple dimensions that defined category membership. Specifically, category membership was defined by both hue and lightness, making these two features jointly predictive of category membership (e.g. a light green object paired with a second light green object). High-density categories varied only in the irrelevant features of size and shape. The low-density categories in the present study contained category members that shared perceptual commonality in a single dimension. Specifically, category membership was defined by lightness only, with category members varying irrelevantly in hue, size and shape (e.g. a light green object paired with a light brown object).

Because higher density and lower density categories pose different obstacles for the learner, their acquisition may be differentially affected by comparison and contrast. Specifically, lower density categories, with members that share less perceptual commonality, may most benefit from comparison learning. Because comparison facilitates attention to the similarities and differences between examples (e.g. Gentner & Markman, 1994), comparison may most benefit category structures with less obvious perceptual similarity. On the other hand, high-density categories may most benefit from contrast learning. Because perceptual similarity is more conspicuous between high-density category members, high-density categories may benefit most from processes that highlight the boundaries between categories.

We investigated how comparison and contrast were mediated by category structure and stimuli features by using a match-to-sample task in which target stimuli were presented with a novel adjective (e.g. 'wug'). To help children to identify that the novel words referred to a property of the object and not to the object name, we used common objects that children were likely to have names for, and we embedded the novel word in an adjectival syntactic frame. Table 1 outlines the design for Studies 1–4. In Studies 1 and 3, the category examples children were presented with did not vary in ways that were irrelevant to the categorization task (i.e. the shape and size of the different exemplars remained constant). In Studies 2 and 4, the category examples varied in a single dimension (i.e. the three examples were either all different sizes or all different shapes). We also varied the category structure. Studies 1 and 2 examined high-density categories, and Studies 3 and 4 examined low-density categories. Because pilot testing indicated that 3-year-olds performed at chance levels when presented with the low-density category structure and 4-year-olds performed near ceiling with the high-density category structure, we presented 3-year-old children with high-density categories and 4-year-old children with low-density categories. Across the four studies, we sought to determine the degree to which the efficacy of comparison and contrast was mediated by the task specifics.

Table 1. Studies 1–4 design

	High-density category structure	Low-density category structure
No variation	Study 1	Studies 3A and 3B
Irrelevant variation	Study 2	Study 4

DOI: 10.1002/icd

STUDY 1

In Study 1, we examined how comparison and contrast processes affected learning of unfamiliar categories. In each trial, children were presented with a category member (e.g. 'This is a wug one') and were asked to select another member of the category (e.g. 'Can you give me another wug one?'). The categories children were presented with were defined by lightness and hue (e.g. dark sienna). One-third of the children were presented with the category in a way that encouraged 'comparison' among three highly similar category members. Another third of the children were presented with the category in a way that encouraged them to contrast a category member with category non-members. We compared performance in these two conditions with a compare–contrast condition in which children compared a category member with another highly similar category member and also contrasted the category member to a category non-member. Because the compare–contrast condition provided both information about category membership and information about category boundaries, this condition was important for benchmarking children's performance when they were provided with optimal category information and provided a control group by which to judge the effectiveness of comparison and contrast alone.

Method

Participants

Participants were 48 3-year-old children ($M = 42.62$ months, $SD = 3.32$ months), 24 boys and 24 girls, randomly assigned and equally distributed across three conditions. In each condition, there was an equal number of girls and boys, and there were no significant differences in children's age between conditions. All participants were recruited and tested in local preschool programs. All children were learning English as their primary language.

Design

Children were randomly assigned to one of the three between-subject comparison conditions: compare, contrast or compare–contrast.

Materials

The stimuli used in this study were three-dimensional, lightweight wooden objects that could be easily handled by the participants. All objects were common objects known to 3-year-old children (e.g. blocks, fish, trees, ducks, butterflies, flowers, hats, balls and boats). The objects ranged in size from 3/4 of an inch to 5 in in length. For sets that manipulated the size of the objects, small objects were 1–2 in in length, medium objects were 2–3 in in length and large objects were 3–5 in in length. When the stimuli presented to children were in a varied size condition, the three objects presented to children were all clearly different from one another in size, and when the stimuli were the same size, they were highly similar in dimension.

Objects were painted in light, medium and dark shades of green, blue and sienna. Paint colors were selected because their original hues were approximately the same value (i.e. lightness) on the Munsell scientific color scale (Munsell, 1913). Paints used were Golden Acrylics in Permanent Green Light (Munsell value = 3.75), Cerulean Blue, Chromium (Munsell value = 4), and Raw Sienna (Munsell value = 4.5). Equal

DOI: 10.1002/icd

parts of original paint colors and neutral grays in light (Golden Acrylics N8 Neutral Gray), medium (Golden Acrylics N5 Neutral Gray) or dark gray (Golden Acrylic N2 Neutral Gray) paints were mixed to produce light, medium and dark shades within each paint hue, altering the value of the original paints without altering their hue. Mixing paint hues with similar original values with the same values of neutral gray and in the same ratio allowed us to create light, medium and dark colors of a similar value across hues.

Exemplar sets. There were six exemplar sets (one set for each of six target properties: light green, dark green, light blue, dark blue, light sienna and dark sienna). Each set contained three identical wooden blocks that did not vary in shape or size. As Figure 1 shows, children in the compare condition were presented with three exemplars from the same lightness category (e.g. three light green cubes). Children in the contrast condition were presented with three exemplars each from a different lightness category (e.g. one light green cube, one medium green cube and one dark green cube). Children in the compare–contrast condition were presented with two exemplars from the same lightness category and one exemplar from a different category (e.g. two light green cubes and one dark green cube).

All children saw all six exemplar sets, such that each child made selections to three sets with a light target (i.e. light green, light blue and light sienna) and three sets with a dark target (i.e. dark green, dark blue and dark sienna). Each hue and lightness combination was labeled with a unique randomly assigned novel adjective (e.g. light green called 'wug' and dark blue called 'fess'). Order of exemplar sets was randomized for each child.

Test sets. Each test set contained three objects from the same hue but each from a different lightness: (e.g. a light green object, a medium green object and a dark green object). Test sets matched the hue of the labeled exemplar sets. Figure 2 shows the four types of test sets: (i) one set that contained objects that were same shaped and same sized (e.g. one light green flower, one medium green flower and one dark green flower, all of the same size); (ii) one set that contained objects that were same shaped and different sized (e.g. one small-sized light green duck, one medium-sized dark green duck and one large-sized medium green duck); (iii) one set that contained objects that were different shaped and same sized (e.g. one light green shark, one medium green tree and one dark green car, all of the same size); and (iv) one set contained objects that were different shaped and different sized (e.g. one small-sized dark green car, one medium-sized light green truck and

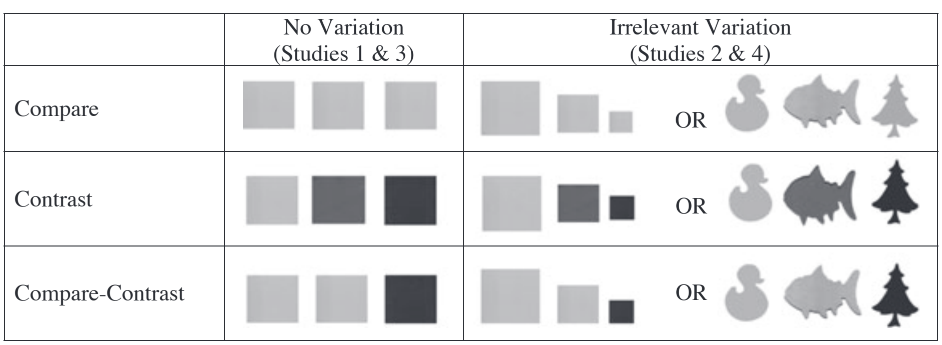

Figure 1. Examples of the labeled stimuli used in Studies 1–4.

DOI: 10.1002/icd

Test Sets	Same shape, same size	
	Same shape, different size	
	Different shape, same size	
	Different shape, different size	

Figure 2. Examples of the test item sets used in Studies 1–4.

one large-sized medium green boat). In exemplar sets that varied by size, objects and color lightness were randomly assigned such that there was no relationship between size and color lightness. Children selected a match for the target exemplar from the test set. Children completed four randomly ordered test trials (one of each of the four types of test sets) with each exemplar set. Children made a total of 24 selections (i.e. from four test sets for each of six exemplar sets).

Procedure

In all three conditions, the experimenter first placed the exemplar set in front of the child and then labeled the set in accordance with condition. In the compare condition, the experimenter indicated that objects in presented stimuli sets were from the same category by labeling the objects with the same novel word, for example, 'This is a wug one (pointing to object 1), this is a wug one (pointing to object 2), this is a wug one (pointing to object 3). These are all wug'. In the contrast condition, the experimenter indicated that objects in presented stimuli sets were from different categories by labeling only the target object with the novel word, for example, 'This is a wug one (pointing to object 1), this is *not* a wug one (pointing to object 2), this is *not* a wug one (pointing to object 3). This is wug, these are *not* wug'. In the compare–contrast condition, the experimenter indicated that objects in presented stimuli sets were from the same and different categories by labeling the category members with the novel word, for example, 'This is a wug one (pointing to object 1), this is a wug one (pointing to object 2), this is *not* a wug one (pointing to object 3). These are wug, this is *not* wug'. Within each condition, the stimuli were always labeled in the same order.

After labeling the stimuli, the objects were left on the table, in view of the child, and the experimenter placed the first test set in front of the child and said, 'Can

DOI: 10.1002/icd

you hand me the wug one of these?' No feedback was provided about the accuracy of children's responses. Regardless of the child's selection, the experimenter provided a neutral response (i.e. 'Thank you.'). The experimenter cleared the test set, relabeled the exemplar set and placed the next test set in front of the child. This was repeated until all four randomly ordered test sets were presented. The experimenter then presented the remaining exemplar sets by using the same procedure.

Results and Discussion

Children's performance in the three conditions is depicted in Figure 3. Comparison of performance to the 33.3% that would be predicted by chance indicated above chance performance in the compare ($t(15) = 2.21$, $p < .05$), contrast ($t(15) = 4.34$, $p < .01$) and compare–contrast ($t(15) = 4.60$, $p < .01$) conditions.

However, performance also varied by condition, $F(2,45) = 3.97$, $p < .05$. Post hoc analysis revealed that participants in the compare condition ($M = .42$ correct, $SD = .17$) performed significantly lower than participants in the compare–contrast condition ($M = .64$ correct, $SD = .26$; Tukey's honestly significant difference (HSD) $p < .05$) and marginally lower than participants in the contrast condition ($M = .61$ correct, $SD = .27$; Tukey's HSD $p = .07$).

Why did children in the contrast condition tend to score higher than children in the comparison condition? One possibility is that contrasting category members against non-members may have highlighted differences between categories. This may have been of particular importance because the lightness categories in this study had relatively small distinctions between them (e.g. light green versus medium green versus dark green). In previous research, contrasting dissimilar objects made the objects seem more dissimilar (Boroditsky, 2007), thus contrasting objects from different categories may have made the three lightness levels seem more distinct from one another. Contrast may have helped to depress attention to 'green-ness' and heighten attention to lightness.

Conversely, performance in the comparison condition was low relative to the compare–contrast and contrast conditions. On the one hand, this finding is surprising because matching objects that are identical in the target dimension for categorization should be a simple task regardless of whether children are presented with objects to compare or contrast. On the other hand, comparison

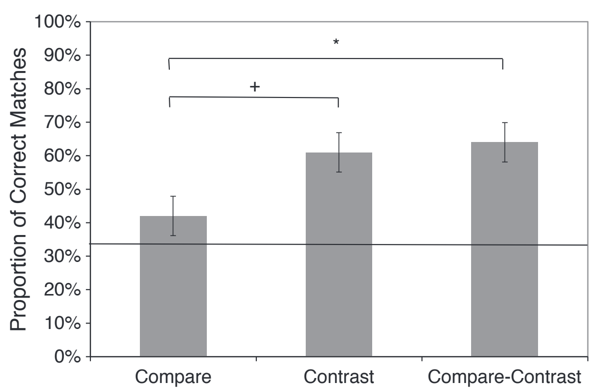

Figure 3. Study 1, mean proportion of correct responses with high-density categories and no variation in irrelevant dimensions by condition (* indicates $p < .05$, + indicates $p < .10$).

DOI: 10.1002/icd

afforded only positive examples of category membership and provided no information regarding category boundaries. Without explicit information regarding non-membership in the category, children may have had no reason to exclude objects that shared any similarity. Similarly, Kovack-Lesh and Oakes (2007) showed that comparison of identical exemplars resulted in children's acquisition of inclusive categories (e.g. containing both dogs and horses), whereas comparison of objects that included irrelevant variation resulted in children's acquisition of exclusive categories (e.g. containing dogs and excluding horses). One possibility is that the children in the current study who compared three identical objects formed a broader category than the children who compared three variable objects. Diminished performance under these conditions is in accord with previous results that find comparison of identical or nearly identical exemplars does not facilitate acquisition with many types of categories (Houston & Jusczyk, 2003; Jamieson & Rvachew, 1992; Lively, Logan, & Pisoni, 1993; Namy, Gentner, & Clepper, 2007; Waxman & Klibanoff, 2000). Comparison of identical exemplars may provide no more information than viewing a single exemplar. Gentner (2003) suggested that comparison is effective because alignment of exemplars highlights common features that are relevant for categorization. When identical exemplars are compared 'all' features become aligned, not just relevant features. Indeed, Namy et al. (2007) found that children who were shown nearly identical examples classified objects on the basis of perceptual similarity; however, children who viewed similar but not identical exemplars classified on the basis of conceptual similarity. Thus, when the exemplars were too perceptually close, comparison appeared to inhibit abstraction.

Notably, performance did not appear to be affected by the particulars of the test sets. A repeated measures ANOVA comparing the four test set types by condition found no differences between the four test sets, $F(3, 43) = 1.19$, not significant (n.s.) and no interaction between test set and condition $F(6, 88) = 1.83$, n.s.. We also examined whether performance changed over the course of the study. However, performance in the first half of the trials versus the second half of the trials did not significantly differ, $t(47) = 1.1$, n.s.

Because children's category acquisition is sensitive to similarity between category members (e.g. Gentner, 2005; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976), we expected that features of the objects that they extended to would also influence category extensions; however, the results did not support this hypothesis. One possibility is that the particulars of the test objects may not have affected performance because the category members shared high perceptual similarity. The features of test objects may have a greater impact when perceptual matches are less transparent. We investigate this possibility in Studies 2–4.

STUDY 2

In Study 2, we replicated the design and procedures of Study 1 by using the same high-density category (i.e. lightness), but we provided category exemplars that varied in features other than lightness (e.g. size or shape). One possibility is that this type of irrelevant feature variation may increase the effects of comparison because experiencing a wider range of category instances provides children with fuller information about category membership than do the highly similar exemplars used in Study 1. On the contrary, irrelevant variation may decrease the effects of contrast. Contrasting objects from different categories highlights category boundaries; however, when objects differ along multiple dimensions, identifying the relevant contrast for categorization becomes more difficult.

DOI: 10.1002/icd

Method

Participants

Participants were 48 3-year-old children ($M = 42.41$ months, $SD = 3.39$ months), 24 boys and 24 girls, randomly assigned to conditions. In each condition, there was an equal number of girls and boys, and there were no significant differences in children's age between conditions. All participants were recruited and tested at local preschool programs.

Design

The design was the same as in Study 1.

Materials

The materials were the same in Study 1 with the addition of exemplar sets that varied in size or shape. Size-varied exemplar sets consisted of three different-sized wooden objects, including one small (.75 × .75 × .75 in), one medium (1 in) and one large (2 in) blocks. The shape-varied exemplar sets consisted of three same-sized wooden objects of different shapes (e.g. a truck, a boat and a car).

Exemplar sets

Unlike Study 1, the exemplar sets included irrelevant variation (i.e. variation in a dimension other than lightness). Three sets included variation in shape, and three sets included variation in size. Figure 1 depicts examples of the exemplar sets.

Test sets

The test object sets were identical to the ones used in Study 1.

Procedure

The procedure was identical to Study 1.

Results and Discussion

Figure 4 shows the performance of children in each condition. Comparison of performance to the 33.3% that would be predicted by chance indicated performance was marginally above chance in the comparison condition ($t(15) = 2.01$, $p = .06$) and significantly above chance in the contrast ($t(15) = 3.21$, $p < .01$) and compare–contrast ($t(15) = 4.04$, $p < .01$) conditions.

Analysis of performance by comparison condition did not find differences between the conditions. An ANOVA indicated that children in the compare ($M = .48$ correct, $SD = .3$), contrast ($M = .53$ correct, $SD = .25$) and compare–contrast ($M = .61$ correct, $SD = .27$) conditions did not significantly differ in extending the target category, $F(2, 45) = .87$, n.s.. Further, an independent samples *t*-test comparing performance in the compare and compare–contrast conditions confirmed that the conditions did not statistically differ, $t(30) = 1.26$, n.s.. In addition, there were no differences in children's performance in the first half of the study versus the second half of the study, $t(47) = .84$, n.s..

We also asked if the dimension of variation affected performance. Because half of the stimuli varied in shape and half of the stimuli varied in size, we examined whether one of the dimensions influenced children's performance more than the other. However, a paired samples *t*-test revealed no differences, $t(44) = .27$, n.s..

DOI: 10.1002/icd

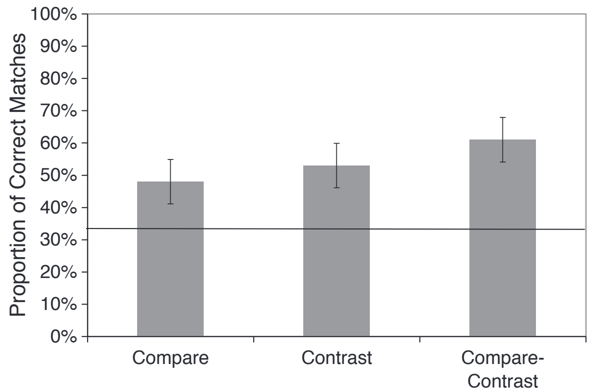

Figure 4. Study 2, mean proportion of correct responses with high-density categories and variation in one irrelevant dimension by condition.

Further, a repeated measures ANOVA comparing test sets and condition revealed that performance did not statistically differ between the four test sets.

The performance of children in the comparison condition did not differ from the performance of children in the compare–contrast condition. Comparison may affect categorization by damping down attention to irrelevant features not shared across category examples and increasing attention to relevant features shared across examples. By experiencing multiple members of the same brightness category that varied in size or shape, attention to size and shape dimensions may have decreased, whereas attention to lightness increased.

Variation also did not appear to negatively affect contrast. One possible reason that irrelevant feature variation did not have a deleterious effect is that the categories children extended were defined by high-density categories with strong perceptual similarity between category members, and the perceptual similarity alone may have aided in binding category members together regardless of irrelevant variation. Low-density categories (e.g. relational categories) that share less perceptual overlap may be more influenced by irrelevant variation. We examine this possibility in Studies 3A, 3B and 4.

The results of Study 2 are consistent with the idea that the amount of variation between category exemplars mediates the efficacy of comparison and contrast. Whereas in Study 1, the comparison condition performed lower than the contrast and compare–contrast conditions, there were no differences in performance across conditions in Study 2. This suggests that one reason that previous research has reached different conclusions regarding the relative effectiveness of comparison and contrast is that the amount of irrelevant variation in the stimuli features has varied between studies.

STUDY 3A

In Studies 1 and 2, we presented children with a category learning task that encouraged comparison or encouraged contrast. However, in these studies, we presented children with high-density categories in which the categories shared cohesive perceptual similarity regardless of irrelevant variation (e.g. all light

DOI: 10.1002/icd

green entities). In Studies 3 and 4, we examined how comparison and contrast affect categorization in lower density categories, specifically light versus dark colors. Because these categories contain greater perceptual variation (e.g. dark categories include dark yellow, dark gray and dark purple), the category as a whole contains less perceptual cohesion, and thus, the presence of irrelevant feature variation in category exemplars may present particular difficulty. Studies 3A and 3B examined comparison and contrast when the exemplars did not vary in irrelevant dimensions (i.e. all exemplars were the same size and shape), and Study 4 examined comparison and contrast when the exemplars provided some irrelevant variation.

Method

Participants

Participants were 60 3-year-old children ($M = 43.47$, $SD = 3.14$), 22 boys and 38 girls, randomly assigned to conditions. Boys and girls were equally distributed across conditions, and there were no significant differences in children's age between conditions. All participants were recruited and tested at local preschool programs.

Design

The design was the same as in Study 1.

Materials

The materials were the same as in Study 1, with one significant change. To present children with a lower density category, the matching object from the test set shared only low perceptual similarity with target stimuli. The matching test set objects matched target objects in color lightness but not in hue. Blue hued stimuli (Golden Acrylic Cerulean Blue, Chromium, Munsell scale value = 4) were matched to green–gold hued test objects (Golden Acrylic Green Gold, Munsell value = 4.25), and sienna hued stimuli (Golden Acrylic Raw Sienna, Munsell value = 4.5) were matched to turquoise hued test objects (Golden Acrylic Cobalt Turquois, Munsell value = 3.75).

Procedure

The procedure was the same as in Study 1; however, there were fewer trials. Participants in Study 3A completed four exemplar sets (two sets of each hue). Children made four matches per stimuli set, for a total of 16 selections. Although no children dropped out of Studies 1 and 2 because of fatigue, the number of trials was decreased for Study 3A because the overall length of the study (approximately 15–20 min) was long for preschool-aged participants. Additionally, comparisons of children's performance in the first versus second half of Studies 1 and 2 revealed no difference, suggesting that the additional trials were not needed to determine the effects of comparison and contrast.

Results and Discussion

As Figure 5 shows, across all conditions, the 3-year-olds' performance was low. Comparison of children's performance to chance (33.33%) showed that the proportion of correct matches in the contrast condition ($M = 38\%$ correct) did not differ from chance and performance in the compare condition ($M = 43\%$ correct;

DOI: 10.1002/icd

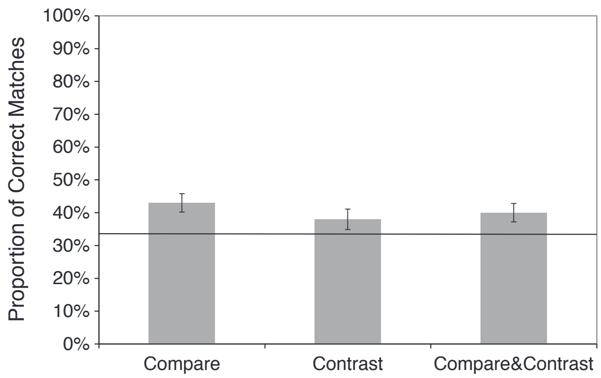

Figure 5. Study 3A, mean proportion of correct responses with low-density categories and no variation in irrelevant dimensions by condition.

$t(19) = 3.24$, $p < .01$), and the compare–contrast condition ($M = 40\%$ correct; $t(19) = 2.3$, $p < .05$) was statistically above what would be expected by chance. However, there was no difference in performance across any of the conditions, $F(2, 57) = 1.5$, $p = .23$. When we adopted the binomial formula (p □ .05) to analyze the performance of individual participants, we found that no children from any condition met criterion for consistently matching on the basis of the target category.

Because members of high-density categories share more salient perceptual similarity than members of low-density categories, acquisition of low-density categories requires learners to look beyond overall perceptual similarity to find subtler, less perceptual, relationships between objects. Overall, the 3-year-olds in Study 3A were not very successful at finding the low-density category match. This is not surprising given the large body of research showing a protracted time course for acquiring low-density categories (e.g. relational categories) versus high-density categories (e.g. object categories) (Clark, 1993; Gentner, 2005; Gentner & Ratterman, 1991; Hall & Waxman, 1993; Keil & Batterman, 1984). In Study 3B, we present 4-year-olds with the same low-density category-matching task.

STUDY 3B

The participants in Studies 3B and 4 were 4-year-old children, in contrast to the 3-year-old children in Studies 1, 2 and 3A. Study 3A revealed that 3-year-old children were unable to perform the low-density categorization task at levels above chance. These results are in alignment with previous research showing that acquiring low-density categories (such as relational categories and categories defined by higher order similarity) poses particular difficulty for young children (Clark, 1993; Gentner, 2005; Gentner & Ratterman, 1991; Hall & Waxman, 1993; Keil & Batterman, 1984; Kloos & Sloutsky, 2008). Because the focus of these studies was on how the particulars of the stimuli affect performance rather than performance at a particular age, we conducted Studies 3B and 4 with 4-year-olds who showed more variation in the low-density category task.

DOI: 10.1002/icd

Method

Participants

Participants were 48 4-year-old children ($M = 53.18$ months, $SD = 3.59$ months), 24 boys and 24 girls, randomly assigned to conditions. In each condition, there were equal numbers of girls and boys, and there were no significant differences in children's age between conditions. All participants were recruited and tested at local preschool programs.

Design

The design was the same as in Study 1.

Materials

The materials were the same as in Study 3A.

Procedure

The procedure was the same as in Study 3A.

Results and Discussion

As Figure 6 shows, comparison of children's performance to chance (33.33%) showed that the proportion of correct matches in the compare ($t(15) = 2.16$, $p < .05$), contrast ($t(15) = 3.24$, $p < .01$) and compare–contrast ($t(15) = 6.04$, $p < .01$) conditions was significantly greater. However, an ANOVA measuring children's performance by comparison condition showed that children's ability to extend lower density categories was significantly affected by whether they compared, contrasted or both compared and contrasted exemplars, $F(2, 45) = 3.73$, $p < .05$. Participants in the compare–contrast condition ($M = .64$ correct, $SD = .2$) performed significantly higher than participants in the compare condition ($M = .45$ correct, $SD = .23$; Tukey's HSD, $p < .05$). There was no difference in performance of children in the contrast condition ($M = .49$ correct, $SD = .2$) and any other condition. Thus, children had relatively less success extending a lower density category when they compared only among similar stimuli.

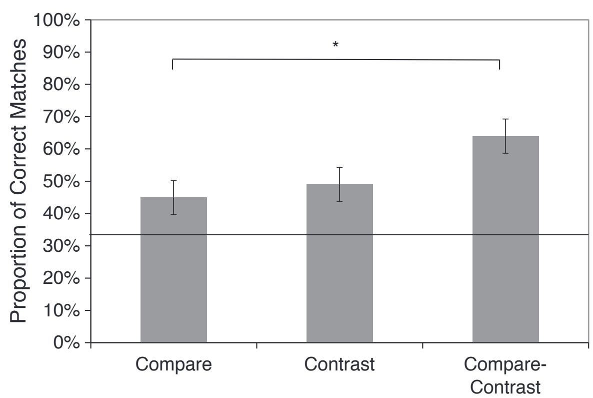

Figure 6. Study 3B, mean proportion of correct responses with low-density categories and no variation in irrelevant dimensions by condition.

DOI: 10.1002/icd

There were no differences in children's performance in the first half of the study versus the second half of the study, $t(47) = 1.35$, n.s.. In addition, there was no effect of size or shape features of test objects; rather, children were equally successful at matching to all four types of test object sets, $F(3, 43) = 1.62$, n.s..

The results of Study 3B followed the same basic pattern as the results of Study 1. As in Study 1, performance was relatively higher in the contrast condition than in the compare condition. This study provides further support for the idea that comparing highly similar or identical exemplars does not strongly benefit categorization. Instead, contrasting exemplars from different categories may be more beneficial. Without variation as a guide to categorization, children may benefit more from explicit information regarding non-membership to the target category.

STUDY 4

Method

Participants

Participants were 48 4-year-old children ($M = 53.71$ months, $SD = 3.51$ months), 24 boys and 24 girls, randomly assigned and equally distributed across conditions. In each condition, there was an equal number of girls and boys, and there were no significant differences in children's age between conditions. All participants were recruited and tested at local preschool programs.

Design

The design was the same as in Study 1.

Materials

The materials were the same as described in Study 3A. In addition, as in Study 2, the exemplar sets included irrelevant variation (i.e. variation in a dimension other than brightness). Two sets included variation in shape, and two sets included variation in size.

Procedure

The procedure was the same as in Study 3A.

Results and Discussion

As Figure 7 shows, analysis of children's performance against chance (33.3%) revealed that performance in the compare ($t(15) = 2.63$, $p < .05$) and compare–contrast ($t(15) = 2.81$, $p < .01$) conditions was significantly above chance; however, performance in the contrast condition ($t(15) = .07$, n.s.) did not differ from chance. Analysis also revealed significant differences across conditions. An ANOVA of children's performance by condition indicated a main effect of condition ($F(2, 45) = 3.88$, $p < .05$). Post hoc analysis indicated that performance in the contrast condition ($M = .33$ correct, $SD = .11$) was significantly lower than performance in the compare–contrast condition ($M = .50$ correct, $SD = .24$) (Tukey's HSD $p < .05$). There was no difference in the performance of children in the compare condition ($M = .42$ correct, $SD = .14$) and any other condition.

DOI: 10.1002/icd

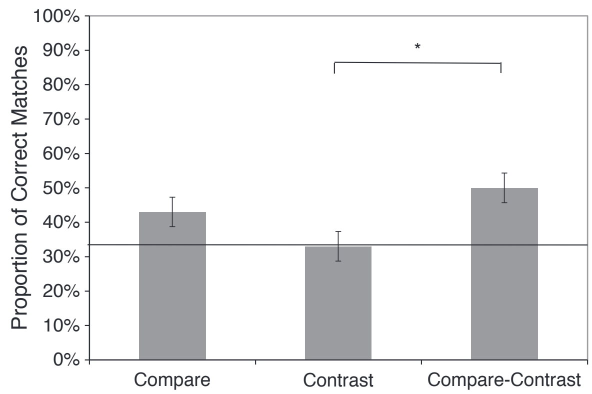

Figure 7. Study 4, mean proportion of correct responses with high-density categories and variation in one irrelevant dimension by condition.

There were no differences in children's performance in the first half of the study versus the second half of the study, $t(47) = .85$, n.s.. We also did not find differences between performance when the variation was in size or shape, $t(47) = .998$, $p > .05$, and there were no differences in performance between the four types of test sets.

Unlike Studies 1, 2 and 3B, contrast was less effective than comparison in promoting category extensions. In fact, across all four studies, contrasting exemplars from low-density categories that varied in an irrelevant dimension emerged as the only condition in which children performed below chance. When categories shared less perceptual similarity, irrelevant variation diminished children's ability to find category matches. Contrast may have been more affected by variation because there was not high perceptual similarity between category members to otherwise aggregate the instances. When category members are less similar in the relevant dimension and vary in irrelevant dimensions, children may require explicit information about category membership that contrast does not provide. High variability between exemplars may have prevented children in the contrast condition from accomplishing a central task of categorization: discarding irrelevant features from a set of possible features shared across examples to determine the relevant dimension for categorization (Sandhofer & Doumas, 2008). When objects did not share many common features, children may have been unable to discard any features as irrelevant for categorization.

Analysis of Data Across Studies 1, 2, 3B and 4

As described in Table 2, children's performance across studies appeared to be influenced by category density and feature variation. A 2 (density) $\times$ 2 (variation)

Table 2. Results of Studies 1, 2, 3B & 4

Study	Age (years)	Category structure	Irrelevant features	Comparison (%)	Contrast (%)	Compare–contrast (%)
1	3	High density	No variation	42	61	64
2	3	High density	Variation	48	53	61
3B	4	Low density	No variation	45	49	64
4	4	Low density	Variation	43	33	50

DOI: 10.1002/icd

ANOVA analyzing the performance of children across the four studies revealed a main effect of category density ($F(1, 188) = 5.12$, $p < .05$) and a main effect of feature variation ($F(1, 188) = 3.67$, $p < .05$). There was no interaction between density and variation, $F(1, 188) = 1.96$, $p > .05$. The average performance of children in Studies 1 and 2 who extended a high-density category ($M = 55\%$ correct, $SE = .02$) was significantly higher than the average performance of children in Studies 3B and 4 who extended a low-density category ($M = 47\%$ correct, $SE = .02$). Importantly, this was the case despite the fact that the children in Studies 3B and 4 were a year older than the children in Studies 1 and 2. Altogether, these results suggest that category density mitigated the efficacy of contrast. The current studies also confirmed that specific features of the category examples affected the relative efficacy of contrast. The average performance of children in Studies 1 and 3 ($M = 54\%$ correct, $SE = .02$), who contrasted exemplars that did not vary in irrelevant dimensions, was significantly higher than the average performance of children in Studies 2 and 4 ($M = 48\%$ correct, $SE = .02$), who contrasted exemplars that varied in one irrelevant dimension. Thus, the results indicate that specific features of the category examples affect the efficacy of contrast.

Figure 8 shows that performance in the comparison (42–48%) and compare–contrast (50%–64%) conditions was somewhat consistent across studies, whereas performance in the contrast condition varied considerably according to specific aspects of the categorization task (33–61%). ANOVA examining children's performance in the comparison conditions ($F(3, 60) = .26$, n.s.) and compare–contrast conditions ($F(3, 60) = 1.17$, n.s.) across the four studies confirmed that there were no differences. Whereas ANOVA examining children's performance in the contrast conditions across the four studies confirmed that there were significant differences, $F(3, 60) = 5$, $p < .01$. Thus, the process of contrasting category members with nonmembers appeared more susceptible to the specifics of the stimuli features and category structure.

GENERAL DISCUSSION

Previous literature has emphasized that viewing multiple instances simultaneously facilitates categorization more than viewing the same instances sequentially

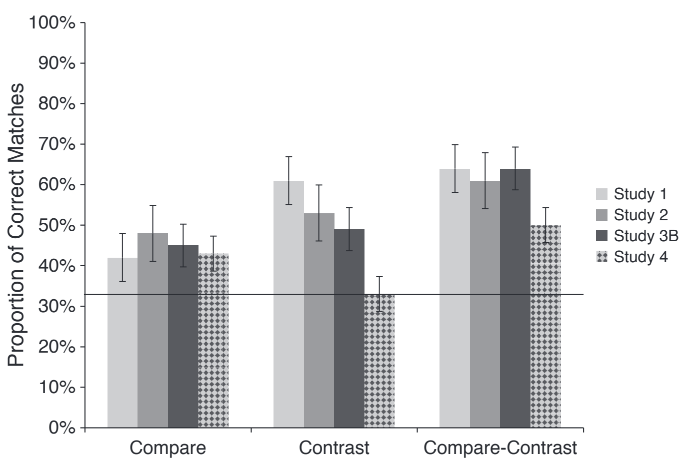

Figure 8. Mean proportion of correct responses by condition across all studies.

DOI: 10.1002/icd

(Anggoro et al., 2005; Boroditsky, 2007; Gentner et al., 2007, 2003; Gentner & Namy, 1999; Kurtz et al., 2001; Loewenstein & Gentner, 2001; Loewenstein et al., 1999; Namy & Gentner, 2002; Namy et al., 1997; Oakes & Ribar, 2005; Rittle-Johnson & Star, 2007). In this literature, the distinction between comparison (viewing multiple examples from the target category) and contrast (viewing a member of the target category with non-members of the category) is seldom drawn. Instead, the extant evidence often conflates comparison and contrast, ignoring significant differences in the two processes. The few studies that have examined the relative effects of comparison and contrast (e.g. Andrews et al., 2005; Kurtz & Boukrina, 2004) have garnered mixed results. We have suggested that one reason is that the specific circumstances surrounding children's use of comparison and contrast may have a substantial influence on the efficacy of these processes. Indeed, in the current studies, we found that category structure and variation in features of the examples interacted with the efficacy of comparison and contrast.

However, comparison and contrast were not equally affected by task specifics. Table 2 illustrates that changes in stimuli features and category structure had the most substantial effect when children contrasted category members against non-members. Why did the performance of children in the comparison condition remain relatively stable across the four studies despite changes in the task specifics? One possibility is that comparison is inherently more constrained than contrast. For example, in learning the category of birds, comparison exemplars are confined only to members of that category (e.g. pigeon, canary and flamingo), whereas the potential contrast exemplars are infinite and varied (e.g. bat, platypus and football). The difference in scope and variability of possible exemplars may influence the efficacy of these processes and their interaction with task specifics. Additionally, the exemplars in the current investigation were restricted, such that comparison exemplars were members of the same narrowly defined category (e.g. light green), and the contrast exemplars were members of a close category (e.g. medium green). Moreover, the exemplars never varied in more than one irrelevant feature, and the type of category structure was introduced only in the extension phase of the trial. Children did not have access to information about whether the novel category was based on hue and lightness (high density) or lightness only (low density) prior to the extension test. This is ecologically sound in some situations as children do not have a priori knowledge of the range of category members in extending real-world categories; however, future research should systematically vary the range of exemplars to determine how comparison and contrast are affected by the range of variation that children may encounter in real-world category learning.

A second possibility for the relatively stable performance in the comparison condition across studies is that the information that comparison provides was equally informative across the four categorization tasks. By comparing multiple members of the target category (e.g. three 'dax' objects), children received positive examples of a category and learned information about the range of category members and necessary features for category membership. Because comparison only provided information about category membership and no information about category non-membership, comparison may have been less affected by task specifics presented in these four studies. On the other hand, contrasting a category member with non-members (e.g. one example that is 'dax' and two that are not) provided children with negative evidence regarding category membership and information regarding category boundaries. Children's dissimilar scores in the contrast conditions across studies suggest that the utility of information about category boundaries may be particularly affected by specifics of the categorization task.

DOI: 10.1002/icd

Feature Specifics

We found that the features of the category instances affected the contrast process. Performance was higher when the examples contained no irrelevant feature variation, and performance was lower when the examples contained irrelevant variation. Similarly, Waxman and colleagues (Klibanoff & Waxman, 2000; Waxman & Klibanoff, 2000; Waxman & Markow, 1998) found that children were successful at extending a novel adjective when they contrasted different colored objects that were the same in basic level kind (e.g. a red toothbrush with a blue toothbrush), but children were not successful when they contrasted objects that were different in basic level kind (e.g. a red toothbrush with a blue plate). That is, categorization was difficult when the example objects contrasted both in the relevant dimension for categorization and contrasted in other ways. However, our results indicate that variation in irrelevant features affected some category structures more than others. For high-density categories with strong perceptual similarity between category members, variation affects categorization less than it does low-density categories.

Why would feature variation be detrimental to the process of contrasting? Contrasting entities that vary in multiple dimensions—including the dimension of interest—may provide too much overall variation to discover the relevant contrast. Thus, the overall similarity between exemplars, both in the target and irrelevant dimensions, proved important for categorization. Detecting the relevant features for categorization may require aligning common features between simultaneously viewed representations (e.g. Gentner, 2005). However, when children viewed objects that were both different in the target dimension as well as in an irrelevant dimension, there may not have been enough common features between the objects to facilitate the type of alignment that leads to effective categorization.

Although manipulations in the features of exemplar sets affected children's performance, manipulations in the features of the test objects had no effect on performance. Across all four studies, there was no difference in performance with the different types of test sets. One possibility is that children may have attended less to the features of the test objects than to the examples that were described using novel adjectives. Previous research shows that words heighten children's attention to objects and facilitate detection of similarities between category members (Gentner, 2005; Loewenstein & Gentner, 2005; McDuffie, Yoder, & Stone, 2006; Smith, 2003; Waxman & Markow, 1995).

Category Structure

Our findings regarding category density support those of Kloos and Sloutsky (2008) showing that high statistical density supports category extension. Overall, children performed better when extending a high-density category than when extending a low-density category. In fact, extension of the low-density category proved impossible for children under 4 years old, requiring the comparison of performance of children of different ages across the two types of categories. Although children's performance was in alignment with previous research documenting a shift in children's ability to acquire categories on the basis of non-obvious or relational information during the preschool years (Blades & Cooke, 1994; Gentner, 1988; Loewenstein & Gentner, 2005) and research documenting the relative difficulty of low-density categories (Kloos & Sloutsky, 2008), future research should tease apart the influences of age and category density.

We found that the contrast process was most strongly affected by category structure. In high-density categories, target stimuli and correct matches shared

DOI: 10.1002/icd

strong perceptual similarity in multiple dimensions. Children received information about category membership from the strong perceptual similarity of target stimuli and correct matches and information about category boundaries from the contrast process. However, in low-density categories, this was not the case as the low-density category members were less perceptually similar, and thus, children received less information about category membership and more information about category boundaries. Information about category boundaries alone did not enable children to extend a low-density category.

The categories acquired in the current studies were based on object properties. Because our results suggest that the effects of comparison and contrast are dependent upon the specific categories being learned and the specific category examples that are available, it is possible that these processes would operate differently in learning other types of categories. For example, comparison and contrast information may be more or less helpful depending on whether a child is extending a novel adjective versus a novel noun. The different developmental trajectory of learning object words versus property words emphasizes the dissimilar challenges that these classes of words may pose for learners. By 2 years of age, children typically exhibit a 'shape bias' in which shape, rather than size, color, or material, dominates the categorization of novel objects (Gershkoff-Stowe & Smith, 2004; Landau, Smith, & Jones, 1988; Smith, 2005). For example, young children are much more successful in their acquisition of object words than in their acquisition of color words (Gottfried & Tonks, 1996; Heibeck & Markman, 1987; O'Hanlon & Roberson, 2007; Thom & Sandhofer, 2009). One possible explanation for this difference may be category density. In general, object categories bound by shape and functional similarity may share more features jointly predictive of category membership and thus higher density than property categories, making category learning easier. Although the current studies shed light on the relationship between category density and the processes of comparison and contrast, it would be important to investigate whether these processes operate differently in acquiring categories from different word classes.

Previous research reflects wide variation in children's acquisition of property words (e.g. color, texture and number) during the preschool years. Children's ability to acquire property words has been linked to factors including the type of language cues provided (Carey & Bartlett, 1978) and children's prior word learning experience in the same domain (Bloom & Wynn, 1997; Soja, 1994; Thom & Sandhofer, 2009). The current studies reveal the particular examples that children experience (i.e. whether they are compared or contrasted, whether they vary in dimensions irrelevant to the target category and whether they are members of dense or sparse categories) as additional factors that influence children's ability to extend property categories.

Our findings suggest that one reason that past research may pose competing accounts of comparison and contrast is that specific aspects of the categorization task impact these processes. Such effects of task specifics have often been noted in studies of adult categorization (e.g. Yeh & Barsalou, 2006), yet little research has examined the effect of the task context on children's categorization (see Kovack-Lesh & Oakes, 2007). Instead, children's word learning is often conceptualized as a simple 1:1 mapping between a word and its referent. However, the current study shows that categorization is not so simple. The specific way that children are familiarized with a category, including whether the objects present are of the same or different categories (comparison or contrast), features of the objects (no variation or irrelevant variation) and category structure (high-density or low-density category) interact to determine categorization. Our results provide

DOI: 10.1002/icd

further evidence that categories are created dynamically during the course of a categorization task (Jones & Smith, 1993; Namy et al., 2007; Schoner & Thelen, 2006; Smith, 2000) and are highly dependent upon task specifics.

ACKNOWLEDGEMENTS

This research was conducted following the relevant ethics guidelines for human research. We would like to thank the families who participated in this research and our dedicated research assistants (Marilyn Corrales, Jenna Edzant, Sarah Gross, Caroline Kalai, Jennifer Menjivar, Naoko Okamoto, Grace Pagunuran and Sanam Salour) for assistance in collecting data.

REFERENCES

Andrews, J. K., Livingston, K. R., & Kurtz, K. (2005). Improving category learning through the use of context items: compare or contrast? In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th annual conference of the cognitive science society* (pp. 121–126). Austin, TX: Cognitive Science Society.

Anggoro, F., Gentner, D., & Klibanoff, R. (2005). How to go from nest to home: children's learning of relational categories. In B. G. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th annual conference of the cognitive science society* (pp. 133–138). Austin, TX: Cognitive Science Society.

Blades, M., & Cooke, Z. (1994). Young children's ability to understand a model as a spatial representation. *Journal of Genetic Psychology*, *155*, 201–218.

Bloom, P., & Wynn, K. (1997). Linguistic cues in the acquisition of number words. *Journal of Child Language*, *24*, 511–533. DOI: 10.1017/S0305000997003188 DOI:dx.doi.org

Boroditsky, L. (2007). Comparison and the development of knowledge. *Cognition*, *102*, 118–128.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the stanford child language conference*, *15*, 17–29.

Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge, England: Cambridge University Press.

Gentner, D. (1988). Metaphor as structure mapping: the relational shift. *Child Development*, *59*, 47–59. DOI: 10.2307/1130388

Gentner, D. (2003). Why we're so smart. In D. Gentner, & S. Goldin-Meadow (Eds). *Language in mind: advances in the study of language and thought*. Massachussetts: MIT Press.

Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe, & D. H. Rakison (Eds.) *Building object categories in developmental time*. Mahway, NJ: Lawrence Erlbaum Associates.

Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: no difference without similarity. *Psychological Science*, *5*, 152–158. DOI: 10.1111/j.1467-9280.1994.tb00652.x

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development*, *14*, 487–513. DOI: 10.1016/S0885-2014(99)00016-7

Gentner, D., & Ratterman, M. J. (1991). Language and the career of similarity. In S. A. Gelman, & J. P. Byrnes (Eds.), *Perspectives on language and thought: interrelations in development* (pp. 225–277). London: Cambridge University Press.

Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development*, *8*, 285–307. DOI: 10.1080/15248370701446434

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: a general role for analogical encoding. *Journal of Educational Psychology, 95*, 393–405. DOI: 10.1037/0022-0663.95.2.393

Gershkoff-Stowe, L., & Smith, L. B. (2004). Shape and the first hundred nouns. *Child Development*, *75*, 1098–1114. DOI: 10.1111/j.1467-8624.2004.00728.x
Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology*, *12*, 306–355. DOI: 10.1016/0010-0285(80)90013-4
Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*, 1–38. DOI: 10.1016/0010-0285(83)90002-6
Gottfried, G. M., & Tonks, S. J. M. (1996). Specifying the relation between novel and known: input affects the acquisition of novel color terms. *Child Development*, *67*, 850–866. DOI: 10.2307/1131865 DOI:dx.doi.org
Hall, D. G., & Waxman, S. R. (1993). Assumptions about word meaning: individuation and basic-level kinds. *Child Development*, *64*, 1550–1570. DOI: 10.2307/1131552
Hampton, J. A., Estes, Z., & Simmons, C. L. (2005). Comparison and contrast in perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1459–1476. DOI: 10.1037/0278-7393.31.6.1459
Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: an examination of fast mapping. *Child Development*, *58*, 1021–1034. DOI: 10.2307/1130543
Houston, D. M., & Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology. Human Perception and Performance*, *29*, 1143–1154. DOI: 10.1037/0096-1523.29.6.1143
Jamieson, D. G., & Rvachew, S. (1992). Remediation of speech production errors with sound identification training. *Journal of Speech Language Pathology & Audiology*, *16*, 519–521.
Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development*, *8*, 113–139. DOI: 10.1016/0885-2014(93)90008-S
Kalish, C. W., & Lawson, C. A. (2007). Negative evidence in inductive generalization. *Thinking and Reasoning*, *13*, 394–425.
Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the development of word meaning. *Journal of Verbal Learning & Verbal Behavior*, *23*, 221–236. DOI: 10.1016/S0022-5371(84)90148-8
Klibanoff, R. S., & Waxman, S. R. (2000). Basic level object categories support the acquisition of novel adjectives: evidence from preschool-aged children. *Child Development*, *71*, 649–659. DOI: 10.1111/1467-8624.00173
Kloos, H., & Sloutsky, V. M. (2008). What's behind different kinds of kinds: effects of statistical density on learning and representation of categories. *Journal of Experimental Psychology. General*, *137*, 52–72. DOI: 10.1037/0096-3445.137.1.52
Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development*, *67*, 2797–2822. DOI: 10.2307/1131753
Kovack-Lesh, K. A., & Oakes, L. M. (2007). Hold your horses: how exposure to different items influences infant categorization. *Journal of Experimental Child Psychology*, *98*, 69–93.
Kurtz, K. J., & Boukrina, O. (2004). Learning relational categories by comparison of paired examples. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th annual conference of the cognitive science society* (pp. 756–761). Mahwah, NJ: Lawrence Erlbaum Associates.
Kurtz, K. J., Miao, C., & Gentner, D. (2001). Learning by analogical bootstrapping. *The Journal of the Learning Sciences*, *10*, 417–446. DOI: 10.1207/S15327809JLS1004new_2
Landau, B., Smith, L. B., & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*, 299–321. DOI: 10.1016/0885-2014(88)90014-7
Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: the role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, *94*, 1242–1255. DOI: 10.1121/1.408177
Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: close comparisons facilitate far mappings. *Journal of Cognition and Development*, *2*, 189–219. DOI: 10.1207/S15327647JCD0202_4
Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology*, *50*, 315–353. DOI: 10.1016/j.cogpsych.2004.09.004
Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review*, *6*, 586–597.

McDuffie, A. S., Yoder, P. J., & Stone, W. L. (2006). Labels increase attention to novel objects in children with autism and comprehension-matched children with typical development. *Autism*, *10*, 288–301. DOI: 10.1177/1362361306063287

Munsell, A. H. (1913). *A Color Notation*. Boston: G. H. Ellis Co.

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of a sow's ears: young children's use of comparison in category learning. *Journal of Experimental Psychology. General*, *131*, 5–15. DOI: 10.1037/0096-3445.131.1.5

Namy, L. L., Gentner, D., & Clepper, L. E. (2007). How close is too close? Alignment and perceptual similarity in children's categorization. *Cognition, Brain, and Behavior, 11*, 647–659.

Namy, L., Smith, L. B., & Gershkoff-Stowe, L. (1997) Young children's discovery of spatial classification. *Cognitive Development*, *12*, 163–184. DOI: 10.1016/S0885-2014(97)90011-3

O'Hanlon, C. G., & Roberson, D. (2007). What constrains children's learning of novel shape terms? *Journal of Experimental Child Psychology*, *97*, 138–148. DOI: 10.1016/j.jecp.2006.12.002

Oakes, L. M., & Ribar, R. J. (2005). A comparison of infants' categorization in paired and successive presentation familiarization tasks. *Infancy*, *7*, 85–98. DOI: 10.1207/s15327078in0701_7

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve problems. *Journal of Educational Psychology*, *99*, 561–574. DOI: 10.1037/0022-0663.99.3.561

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.

Sandhofer, C. M., & Doumas, L. A. A. (2008). Order of presentation effects in learning color categories. *Journal of Cognition and Development*, *9*(2), 194–221. DOI: 10.1080/15248370802022639

Schoner, G., & Thelen, E. (2006). Using dynamic field theory to rethink infant habituation. *Psychological Review*, *113*, 273–299. DOI: 10.1037/0033-295X.113.2.273

Smith, L. B. (2000). From knowledge to knowing: real progress in the study of infant categorization. *Infancy*, *1*, 91–97. DOI: 10.1207/S15327078IN0101_08

Smith, L. B. (2003). Learning to recognize objects. *Psychological Science*, *14*, 244–250. DOI: 10.1111/1467-9280.03439

Smith, L. B. (2005) Shape: a developmental product. In L. Carlson, & E. VanderZee (Eds). *Functional features in language and space* (pp. 235–255). Oxford University Press.

Soja, N. N. (1994). Young children's concept of color and its relation to the acquisition of color words. *Child Development*, *65*, 918–937. DOI: 10.2307/1131428

Thom, E. E., & Sandhofer, C. M. (2009). More is more: the connection between vocabulary size and word extension. *Journal of Experimental Child Psychology*, *104*, 466–473. DOI: 10.1016/j.jecp.2009.07.004

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology*, *36*, 571–581. DOI: 10.1037/0012-1649.36.5.571

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. DOI: 10.1006/cogp.1995.1016

Waxman, S. R., & Markow, D. B. (1998). Object properties and object kind: twenty-one-month-old infants' extension of novel adjectives. *Child Development*, *69*, 1313–1329. DOI: 10.2307/1132268

Yeh, W., & Barsalou, L. W. (2006). The situated nature of concepts. *The American Journal of Psychology*, *119*, 349–384. DOI: 10.2307/20445349