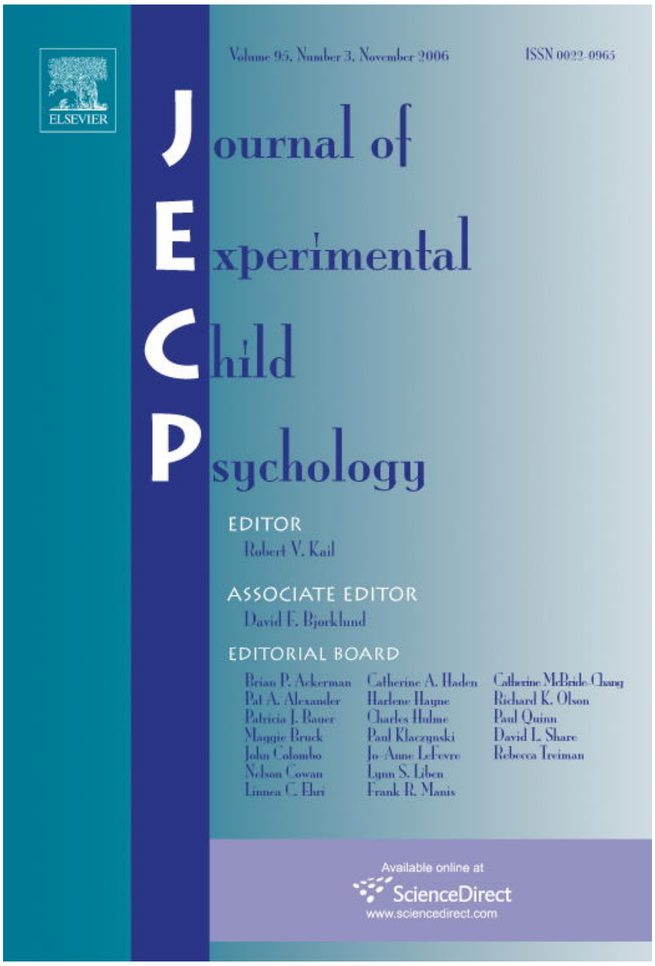
Volume 95, Number 3, November 2006
ISSN 0022-0965
ELSEVIER
Journal of
Experimental
Child
Psychology
EDITOR
Robert V. Kail
ASSOCIATE EDITOR
David F. Bjorklund
EDITORIAL BOARD
Brian P. Ackerman
Pat A. Alexander
Patricia J. Bauer
Maggie Bruck
John Colombo
Nelson Cowan
Linnea C. Ehri
Catherine A. Haden
Harlene Hayne
Charles Hulme
Paul Klaczynski
Jo-Anne LeFevre
Lynn S. Liben
Frank R. Manis
Catherine McBride-Chang
Richard K. Olson
Paul Quinn
David L. Share
Rebecca Treiman
Available online at
ScienceDirect
www.sciencedirect.com

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Journal of Experimental Child Psychology 95 (2006) 194–214
Journal of Experimental Child Psychology
www.elsevier.com/locate/jecp

Preschoolers' use of surface similarity in object comparisons: Taking context into account

Jae H. Paik [a,*], Kelly S. Mix [b]

[a] *Department of Psychology, San Francisco State University, San Francisco, CA 94132, USA*
[b] *College of Education, Michigan State University, East Lansing, MI 48824, USA*

Received 3 November 2005; revised 9 June 2006
Available online 9 August 2006

Abstract

Previous research has emphasized the role of within-match similarity in children's comparisons. The current study investigated another potentially important contributing factor, namely the distinctiveness of the matching items relative to other items in the scene. Using a well-known relational mapping task, we found that 3- and 4-year-olds made more correct matches between identical objects when those objects were maximally distinctive from the foils. In a cross-mapping experiment, where same relative size was pitted against object similarity, 3-year-olds made more incorrect object matches when the objects were both similar to each other and distinctive from the foils. Furthermore, 3- and 4-year-olds performed the same, regardless of within-match similarity, so long as the ratios of within-match and nonmatch similarity were roughly equal. These findings suggest that children's comparisons are guided by a ratio consisting of many pairwise similarity relations, including (but not limited to) the degree of within-match similarity.

Keywords: Similarity; Comparison; Context effects; Cognitive development; Preschool children

Introduction

Research on children's similarity judgments has focused primarily on the relations between the matching objects themselves. This research has demonstrated that early

* Corresponding author. Fax: +1 415 338 2398.
E-mail address: jaepaik@sfsu.edu (J.H. Paik).

doi:10.1016/j.jecp.2006.06.002

comparisons differ from the comparisons of adults in fundamental ways. Rather than picking out specific commonalities, young children tend to base their comparisons on global similarity (Smith, 1984). And rather than matching items that share an underlying relation, young children tend to overemphasize the way things look (DeLoache, 1989, 1990; Gentner & Rattermann, 1991; Gentner & Toupin, 1986; Mandler & Bauer, 1988). This initial reliance on overall or high-similarity matches has been demonstrated time and again and has played a major role in theories of conceptual development.

Although children's comparisons undoubtedly are affected by the similarity between matching items, there is another potentially important factor to consider, namely the contrast between the matching items and other available objects. Items are more likely to be compared, and eventually perceived to be similar, when two conditions are met: (a) the amount of overlap between the items is high (e.g., two red cars) and (b) the amount of overlap between these items and others in the scene is low (e.g., two red cars in a pile of teddy bears). When items overlap significantly with nonmatching items (e.g., two red cars in a pile of fire engines), the perceived similarity of the "matching" items should be low even though the number of features shared between the two items remains high.

The effects of distinctiveness are well established in the literature on adults' similarity judgments. However, these effects have not received as much attention in the developmental literature. This is unfortunate because distinctiveness could play a pivotal role in the earliest emergence of grouping and categories. The current study was aimed at addressing this gap to establish whether children, like adults, use the entire context when noticing and evaluating similarity.

Role of overall similarity in conceptual development

Comparisons are considered both the product and the process of conceptual development and learning. That is, people are thought to (a) compare items because they notice dimensions of similarity (product) and (b) discover new dimensions of similarity by making comparisons (process). But what starts this bootstrapping mechanism in motion? Early in development, children may recognize few, if any, points of comparison. This means that they would have few reasons to compare objects and, therefore, would have limited opportunities to discover new kinds of similarity.

Children's consistent reliance on overall similarity suggests a possible explanation, namely that children may rely on overall similarity because this increases the likelihood of recognizing at least one commonality (Gentner & Medina, 1998; Gentner, Rattermann, Markman, & Kotovsky, 1995; Kotovsky & Gentner, 1996; Smith, 1989, 1993). In other words, the more overlapping features, the more opportunities children have to initiate comparisons. Once initiated, these comparisons could support the discovery of new dimensions of similarity. The more dimensions children learn to recognize, the more likely they will compare less similar objects in the future, thereby discovering additional, less obvious commonalities. In this way, children can move from few to many known attributes or relations, from holistic to specific comparisons and groupings.

To illustrate, consider how children might learn to match objects by color. It might seem that when color is the only shared feature (e.g., red car–red apple), it would automatically increase in salience, leading to a match. However, unless children already understand the relation "same color," they might see no reason to compare these items in the first place. In contrast, when color is one of many shared features (e.g., red car–red truck), children could compare the items after noticing any one of a long list of commonalities, including shape, material, movement, size, and function. Once initiated, the comparison process could help children to discover other dimensions of similarity, such as same color, that they might not have noticed before.

This line of reasoning leads to an interesting prediction. Because children would be more likely to see the similarity between objects with many shared features, the richer and more complex an object is, the greater its perceived similarity to another object should be. In other words, because rich objects have more alignable features than do simple objects, rich objects should be compared more readily, all else being equal (Gentner & Gunn, 2001; Gentner & Markman, 1994; Gentner & Rattermann, 1991; Markman & Gentner, 1993; Tversky, 1977). The developmental implication is that rich objects should play an important role in emerging categories because they would be particularly likely to promote early comparisons.

Support for this prediction was obtained in a study of preschool children's relational mapping ability (Gentner & Rattermann, 1991). In this study, 3- and 4-year-olds were shown identical rows of objects arranged by size (Fig. 1A). In the rich condition, the arrays consisted of complex objects such as a toy house, a toy car, and a decorative mug. In the

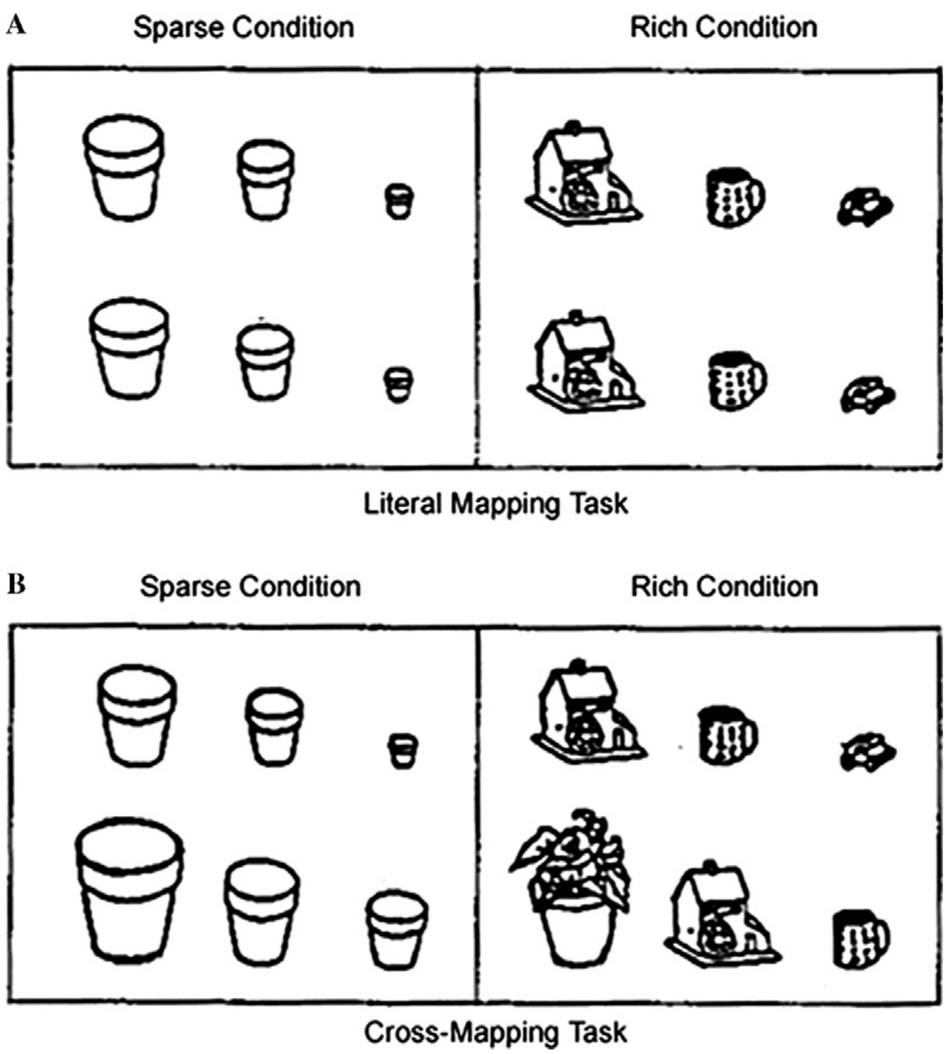

Fig. 1. Sample stimuli used in the relational mapping studies: (A) literal mapping task; (B) cross-mapping task (Gentner & Rattermann, 1991).

sparse condition, the arrays consisted of simple objects such as three clay pots. While children watched, an experimenter hid a sticker under one of three objects in her array. The children's task was to find a sticker under the same object in the other array of objects. Because finding the hidden sticker required nothing more than a straightforward identity match, one might expect performance to be the same for rich and sparse items. However, Gentner and Rattermann (1991) found that both 3- and 4-year-olds performed significantly better when the identical items were rich rather than sparse.

Further evidence of this richness effect was found in another version of the task, where relational similarity was pitted against surface similarity (Fig. 1B).[1] In this condition, children needed to ignore the identical object match and find the sticker under the object with the same relative size (e.g., the largest). Given preschoolers' focus on surface features, it was not surprising that children performed poorly in the cross-mapping task. Interestingly, it was found that children performed less accurately when the objects were rich than when they were sparse. That is, the cross-mapping effect was stronger when the objects were complex and detailed. Gentner and Rattermann concluded that children were pulled even more strongly toward the object match when the objects were rich because having more shared features provides more invitations to compare them.

Shifting context and the development of comparisons

Research on adults' comparisons has shown that similarity judgments depend on more than the features of the matching items themselves. Instead, similarity is based on the ratio of (a) the similarity between the two matching items and (b) the dissimilarity between the matching items and the alternatives (e.g., Barsalou, 1982; Goldstone, Medin, & Halberstadt, 1997; Markman & Gentner, 1993; Rips & Collins, 1993; Sjöberg, 1972; Tversky, 1977). For example, it has been argued that basic level categories are more salient than subordinate and superordinate categories because members of basic level categories (e.g., chair) not only share many perceptual and functional attributes but also differ significantly from members of contrasting categories (e.g., table) (Rosch, 1975). On this view, the dissimilarity between categories serves to highlight the shared attributes within the category, thereby increasing the perceived similarity between category members.

This ratio also seems to underlie shifts in adults' similarity judgments. Several studies have shown that perceived similarity can change, even for the same pair of items, depending on the other items in the scene (Barsalou, 1982; Gati & Tversky, 1984; Goldstone et al., 1997; Rips & Collins, 1993; Sjöberg, 1972; Tversky, 1977). For example, adults judged Sweden to be more similar to Austria than to Hungary when Poland was included in the comparison (Tversky, 1977). However, when Poland was replaced with Norway, Hungary was judged to be more similar to Austria than to Sweden. The grouping of countries was based on political status when Poland was included, whereas it was based on geographical proximity when Norway was included. These changes in similarity judgments reflect the

[1] In the current study and in the original study (Gentner & Rattermann, 1991), spatial position was confounded with relative size. That is, the correct relative size match always was presented directly across from its analogue in the experimenter's array. Paik and Mix (2006) contrasted these two possibilities (spatial position vs. relative size) and found that 5-year-olds' cross-mapping performance declined significantly when the objects were presented in random clusters rather than in parallel lines. Thus, it is possible that when children are correct, they may use spatial position, instead of relative size, as the basis for their comparisons.

context in which countries were grouped. As the context shifts, adults' comparisons adapt flexibly.

These context effects on adults' similarity judgments raise two questions regarding the development of comparisons in young children. First, are children's initial comparisons promoted by variations in distinctiveness in addition to variations in within-match similarity (including richness vs. sparseness)? If so, then the surrounding context may serve to prompt children to make comparisons especially at early stages of development. Second, what is the developmental trajectory of flexible comparisons? If sensitivity to ratios of similarity is what underlies this flexibility in adults, then it is important to know when and how children become sensitive to these ratios. One could argue that children use overall similarity as a default because they are insensitive to changes in context, because they are unable to adapt their responses, or both. Alternatively, they might well be sensitive to them, but this simply has not been tested before. In fact, related research indicates that children are sensitive to a range of contextual factors such as differences in the experimenter's instructions, use of verbal labels, stimulus/problem presentation, response time allowed, and other task demands (e.g., Deák, Ray, & Pick, 2004, 2004; Jones, Smith, & Landau, 1991; Kemler Nelson, Frankenfield, Morris, & Blair, 2000). For example, 3-year-olds usually extend count nouns to new objects that are the same shape. So, if they are taught that a round object is called a "dax," they will choose another round object when asked to find another dax. However, children's choices shift from shape to other properties, such as texture and color, when the named objects have eyes (Jones et al., 1991) or no longer are rigid (e.g., Soja, 1992; Soja, Carey, & Spelke, 1991). Because novel word extensions reflect implicit grouping, these studies suggest that young children's explicit matches would be context sensitive and flexible.

In the current study, we tested directly whether young children's similarity judgments are affected by changes in context. We built on Gentner and Rattermann's (1991) evidence for richness effects. Their relational mapping task provides a straightforward format for manipulating the pairwise similarities among all of the items in the scene, including (but not limited to) the matching items themselves. This permitted the systematic variation of several dimensions of similarity that allowed us to say not only whether context matters but also in what specific ways it does. This task also approximates the conditions under which children likely notice similarity in their everyday play (i.e., similar items surrounded by other items that vary in their similarity to each other).

Experiment 1

Gentner and Rattermann (1991) found that children made more object matches when the objects were rich and complex. However, it is possible that these effects were mediated by the surrounding objects (i.e., the foils). In fact, it is unclear whether there were richness effects at all because richness and distinctiveness covaried in the original study. That is, the rich objects not only were complex and detailed but also were completely different from the other objects in the arrays. In contrast, the sparse objects not only were plain but also were essentially the same as the other objects except for their size. We addressed this problem by varying richness and distinctiveness independently. In Experiment 1, we tested whether children's performance on the identity-mapping task improves as the degree of dissimilarity between objects in the array increases.

Method

Participants

A total of 48 3-year-olds (mean age = 3 years 5 months, range = 3 years 0 months to 3 years 11 months) and 47 4-year-olds (mean age = 4 years 5 months, range = 4 years 0 months to 4 years 11 months) participated. An additional 3 children were excluded from the study because they refused to complete the task. These are the same age groups used in the original relational mapping study (Gentner & Rattermann, 1991). Each age group included approximately the same number of boys and girls. Children were recruited from local preschools or through newspaper advertisements and local birth announcements. All came from the same predominantly white middle-class population and spoke English as their primary language.

Design

Children were randomly assigned to one of four conditions that result from crossing two dimensions: (a) richness of the objects in the two rows (rich vs. sparse objects) and (b) distinctiveness of the objects within row (distinctive vs. similar objects) (Fig. 2). Roughly the same number of children participated in each of the four conditions.

Materials

Objects used in the rich and sparse conditions were the same type, but rich objects were more detailed and elaborate. For example, beds in the sparse conditions were constructed from wooden blocks and painted in a single solid color. However, beds in the rich conditions had additional features such as colorful fabric blankets, pillows, and bed skirts.

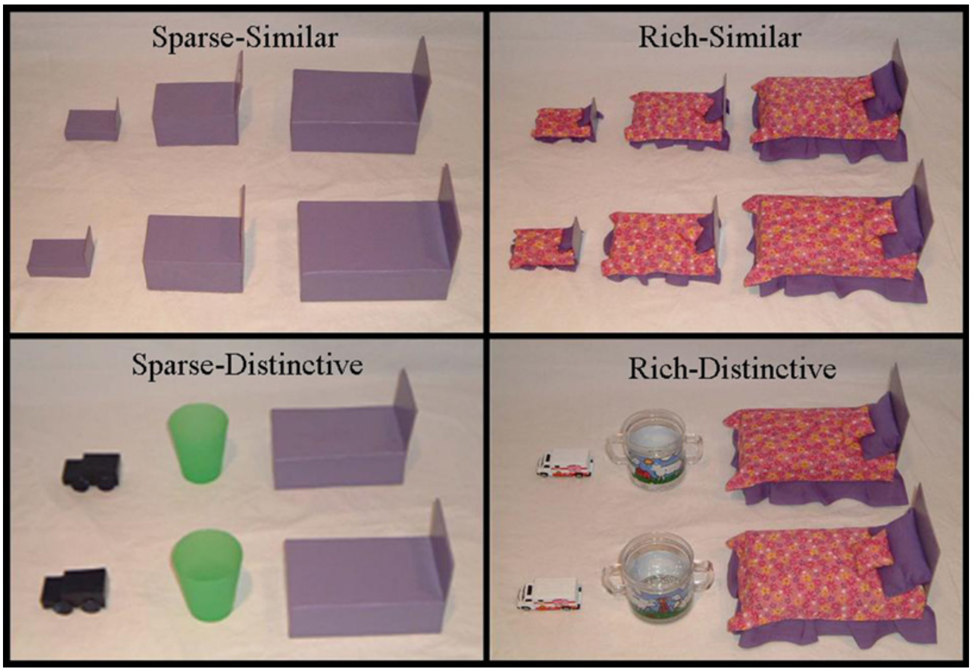

Fig. 2. Sample stimuli conditions in Experiment 1.

Within each condition, the objects varied in distinctiveness relative to the other objects in the array. In the distinctive objects conditions, the three objects were completely different from one another (e.g., a small truck, a medium sofa, and a large bed). In the similar objects conditions, objects differed only in size (e.g., a small bed, a medium bed, and a large bed). (For a complete list of stimuli, see Fig. 2 and the Appendix.) Two distinctive object sets were used. Furthermore, two identical small cups and medium saucers were used as familiarization objects. The dimensions of the different sizes were the same across conditions: $8 \times 5 \times 5$ cm (small), $13 \times 10 \times 10$ cm (medium), and $21 \times 12 \times 10$ cm (large).

Thus, the rich–distinctive and sparse–similar conditions provide a direct replication of the design used previously by Gentner and Rattermann (1991), where richness and distinctiveness were confounded. The rich–similar and sparse–distinctive conditions, unique to the current study, indicated whether distinctiveness has an impact either alone or in combination with richness.

Procedure

Children completed a mapping task similar to that used by Gentner and Rattermann (1991). In this task, the experimenter and the children sat across from each other with two rows of three objects between them (Fig. 2). Within each row, the objects were arranged from smallest to largest, running left to right from the children's point of view. The two rows were 8 in. apart, so that the experimenter's row was clearly separate from the children's row. The row of objects in front of the children was identified as the children's "toys," and the row of objects in front of the experimenter was identified as the experimenter's toys.

Prior to the test trials, two familiarization trials were presented in which both the experimenter and the children had a small cup and a medium saucer arranged in opposite arrays. At the beginning of each trial, the children were asked to hide their eyes behind a cardboard screen (20×14 in.) while the experimenter hid a sticker underneath one of the children's toys. The experimenter asked the children to lower the cardboard and explained, "Now, I'm going to hide *my* sticker under one of my toys." She hid the sticker underneath one of her objects, in full view of the children, and said, "If I hide my sticker under this toy, where do you think your sticker is under your toys?" The children indicated their choice by pointing. When the children correctly identified the object with the sticker, the experimenter gave them the sticker and said, "Good job!" or "You're right!" When the children did not identify the correct object, the experimenter explained, "No, if I hide my sticker under this toy, it would be under this toy in your set" while pointing to the two objects where the stickers were hidden. In one trial the stickers were hidden under the saucers, and in the other trial the stickers were hidden under the cups.

Following the familiarization trials, children completed 6 test trials per object set for a total of 12 mapping trials. Children received feedback in that they were allowed to keep the stickers if they selected the correct objects. If they responded incorrectly, the experimenter silently removed the hidden stickers while the children watched. The experimenter offered no verbal feedback on their performance. These test trials were presented in one of two fixed random orders. The position of the hidden sticker was counterbalanced across trials. Children mostly pointed to and sometimes lifted up the objects they chose. The experimenter recorded only children's first response in coding the data. The entire task, including familiarization, required approximately 20 min to complete.

Results and discussion

Adults' similarity ratings

To ensure that our assumptions about the objects in each condition were correct (i.e., that they varied in similarity and distinctiveness across conditions), similarity ratings were obtained from a group of 15 undergraduate and graduate students. Participants viewed the stimuli in pairs presented in a random order. As each pair was presented, participants rated the similarity between the two objects on a scale of 1 to 10 (1 = *not at all similar to one another*, 10 = *extremely similar to one another*). Participants saw each pair once, resulting in a total of 48 trials (i.e., 6 possible stimuli pairs × 4 conditions × 2 object sets). To facilitate presentation of these 48 pairs, 5 × 8-in. color photographs of the objects were used instead of the objects themselves. However, when children were tested, the actual objects were used.

First, we confirmed that the similarity between target items (i.e., within-match similarity) was the same regardless of distinctiveness (distinctive objects: $M = 9.78$, $SD = 0.19$; similar objects: $M = 9.72$, $SD = 0.18$), $t(11) = 0.72$. Next, we confirmed that the similarity of items within each array (i.e., nonmatch similarity) was lower in the distinctive objects conditions ($M = 2.76$, $SD = 0.60$) than in the similar objects conditions ($M = 8.40$, $SD = 0.28$), $t(11) = 29.44$, $p < .001$. This demonstrates that items in the distinctive objects conditions really were distinctive. An unexpected result was that the similarity ratings did not differ for rich and sparse objects (rich objects: $M = 7.75$, $SD = 3.11$; sparse objects: $M = 7.78$, $SD = 2.80$), $t(23) = 0.27$. This is surprising given the hypothesis that rich items are perceived as more similar than sparse objects. As we will see, however, this pattern was borne out in the developmental data as well. Finally, we confirmed that there was no difference in perceived similarity for the two object sets used in this experiment (Object Set 1: $M = 7.70$, $SD = 2.92$; Object Set 2: $M = 7.60$, $SD = 3.02$), $t(23) = 0.07$.

Children's mapping performance

All children finished the familiarization trials. Fig. 3 shows the mean number of correct responses for each age group in the four conditions. Because this task was essentially an identity match, it is not surprising that children generally performed above chance. However, rather than consistently performing at ceiling in this straightforward comparison, children's performance differed across conditions in revealing ways. First, as in previous research, children made more relational matches in the rich–distinctive condition ($M = 9.83$, $SD = 2.81$) than in the sparse–similar condition ($M = 7.78$, $SD = 3.42$). Thus, we replicated the original effect previously attributed to richness. However, recall that these conditions confound richness with distinctiveness. When these variables were separated (i.e., in the rich–similar and sparse–distinctive conditions), there was evidence that distinctiveness, rather than richness, drove this effect. Specifically, children made more correct matches in the distinctive objects conditions than in the similar objects conditions. Moreover, although 3-year-olds performed above chance in both of the distinctive objects conditions, they performed randomly in the rich–similar condition. Possibly, low distinctiveness had a considerable impact on performance, interfering with children's ability to match even identical objects. This pattern also raises questions about the role of

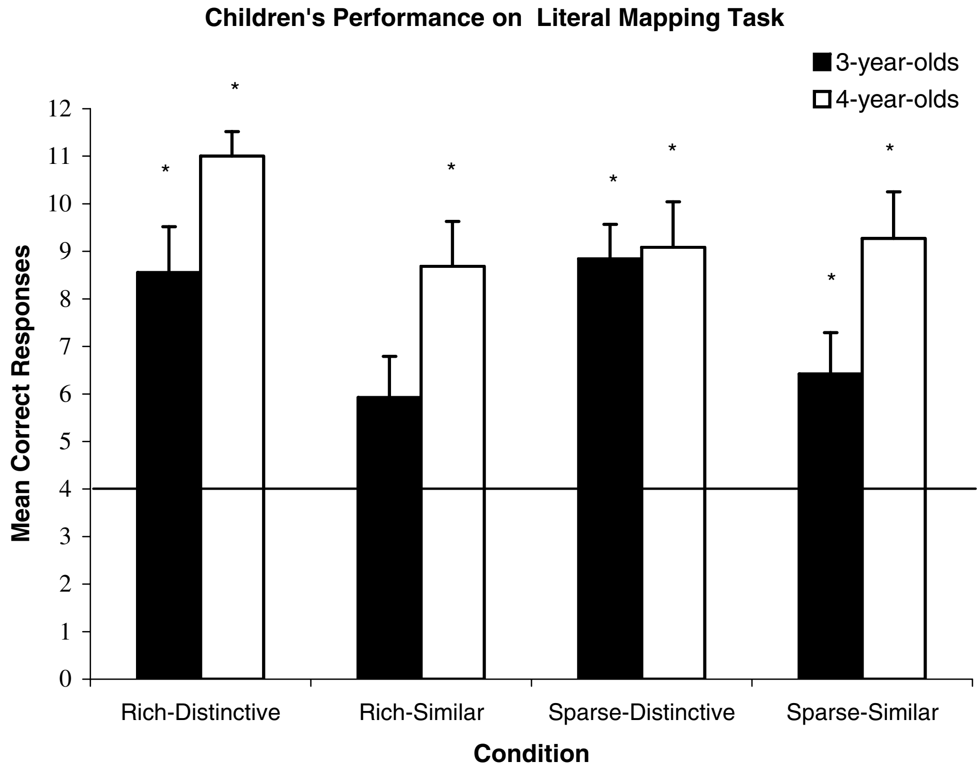

Fig. 3. Children's mean correct responses for the different stimuli conditions in Experiment 1. *Above chance, $ps < .08$.

richness in this task because children performed less accurately when the less distinctive objects were rich than when they were sparse.

An analysis of variance (ANOVA) with age (3- vs. 4-year-olds), distinctiveness (distinctive vs. similar), and richness (rich vs. sparse) as between-participants factors confirmed that the main effect of distinctiveness was significant, $F(1, 95) = 9.20$, $p < .003$. The ANOVA also revealed a significant effect of age, $F(1, 95) = 11.87$, $p < .001$. Specifically, 4-year-olds ($M = 9.51$, $SD = 3.00$) outperformed 3-year-olds ($M = 7.44$, $SD = 3.15$) overall. However, no evidence of a richness advantage was obtained, $F(1, 95) = 0.01$. Furthermore, we found no significant interactions among richness, distinctiveness, and age. This suggests that even though children improved overall with development, they did not surmount the difficulty of the similar objects conditions—at least not in this age range. Note that because scores on this task were proportional, we performed parallel analyses using arcsin transformations of children's scores. These analyses yielded the same basic patterns as those reported previously using the untransformed data.

This finding seems to suggest that the number of alignable features between matching objects does not affect children's comparisons as claimed previously (Gentner & Rattermann, 1991). However, this is not necessarily the case. The problem is that richness effects may be washed out in the ratio of within-match similarity to nonmatch similarity. When all of the items in a scene are rich, there are many points of alignment between matching items. But there are many points of alignment between the nonmatching pairs as well. This means that the ratios for rich and sparse items could end up being equal. Thus, simply having more points of alignment might not be enough to promote a comparison. Instead, the way the points align (i.e., having more similarity within than without) may be what counts.

Another mediating factor that may have obscured the richness effect is that all of the matching objects shared a high degree of surface similarity. They were, in fact, identical. This means that within-match similarity always was at ceiling, leaving the ratio of similarity to vary depending on distinctiveness (i.e., nonmatch similarity) alone. It is an open question whether richness effects might emerge when within-match similarity is more variable. Perhaps richness would make a difference when there is less featural overlap and children need to work harder to find a way to align items.

Experiment 2

Recall that Gentner and Rattermann (1991) also obtained richness effects in another version of the relational mapping task, one in which the relation of size was pitted against object identity. Under those conditions, children performed less accurately when the objects were rich, apparently because they were more drawn to the object match when there were multiple points of alignment. In our version of this cross-mapping task, we tested whether these effects were due to richness, distinctiveness, or both.

Method

Participants

A total of 113 children participated: 56 3-year-olds ($M = 3$ years 6 months, range = 3 years 0 months to 3 years 11 months) and 57 4-year-olds ($M = 4$ years 6 months, range = 4 years 0 months to 4 years 11 months). Approximately the same number of girls and boys were included in each age group. An additional 4 children were excluded from the study because they refused to complete the task. Children were recruited from a predominantly white middle-class population through newspaper advertisements and local birth announcements. All spoke English as their primary language.

Design

As in Experiment 1, children were randomly assigned to one of four conditions: rich–distinctive, rich–similar, sparse–distinctive, or sparse–similar.

Materials and conditions

The three object sizes from Experiment 1 (i.e., small, medium and large) were used. On some trials, extra large objects ($25 \times 18 \times 12$ cm) also were used. For low-distinctiveness conditions (i.e., rich–similar and sparse–similar), the objects were identical except for size. In the high–distinctiveness conditions (i.e., rich–distinctive and sparse–distinctive), the objects differed along a number of dimensions, including size (e.g., an extra large house, a large bed, a medium sofa, and a small truck). As before, the rich and sparse objects were essentially the same except that rich objects were more detailed and elaborate. The familiarization objects were simple blue cubes (small = 5 cm^3, medium = 7 cm^3, and large = 9.5 cm^3). (For a complete list of stimuli, see the Appendix A)

Procedure

We used the same procedure as in Experiment 1 except that the test objects were arranged so that object similarity was pitted against relational similarity (i.e., relative size) (Fig. 4). For example, if the experimenter hid the sticker under her largest object (e.g., a large bed), children had to find the sticker under their largest object (e.g., an extra large house) and ignore the identical object (e.g., a large bed). This was conveyed to children through two familiarization trials that involved comparisons between pairs of different-sized cubes. For example, if the experimenter had a small cube and a medium cube and the children had a medium cube and a large cube, the experimenter hid a sticker under her small cube. Children were supposed to find a sticker under their smaller object, that is, under their medium cube. In the other familiarization trial, the experimenter had a medium cube and a large cube and the children had a small cube and a medium cube. As before, children were given verbal feedback on their familiarization performance but not on the test trials.

Results and discussion

Adults' similarity ratings

A group of 15 undergraduate and graduate students were asked to provide similarity ratings on 72 pairs of objects (9 pairs $\times$ 4 conditions $\times$ 2 object sets). When distinctiveness was low (i.e., similar objects conditions), both the within-match and nonmatch similarity ratings were high and equally so (within-match: $M = 8.48$, $SD = 0.26$; nonmatch: $M = 8.71$, $SD = 0.82$), $t(17) = 0.92$. When distinctiveness was high (i.e., distinctive objects conditions), both the within-match and nonmatch similarity ratings were relatively low (within-match: $M = 2.53$, $SD = 0.80$; nonmatch: $M = 5.22$, $SD = 3.59$). Still, nonmatch similarity ratings

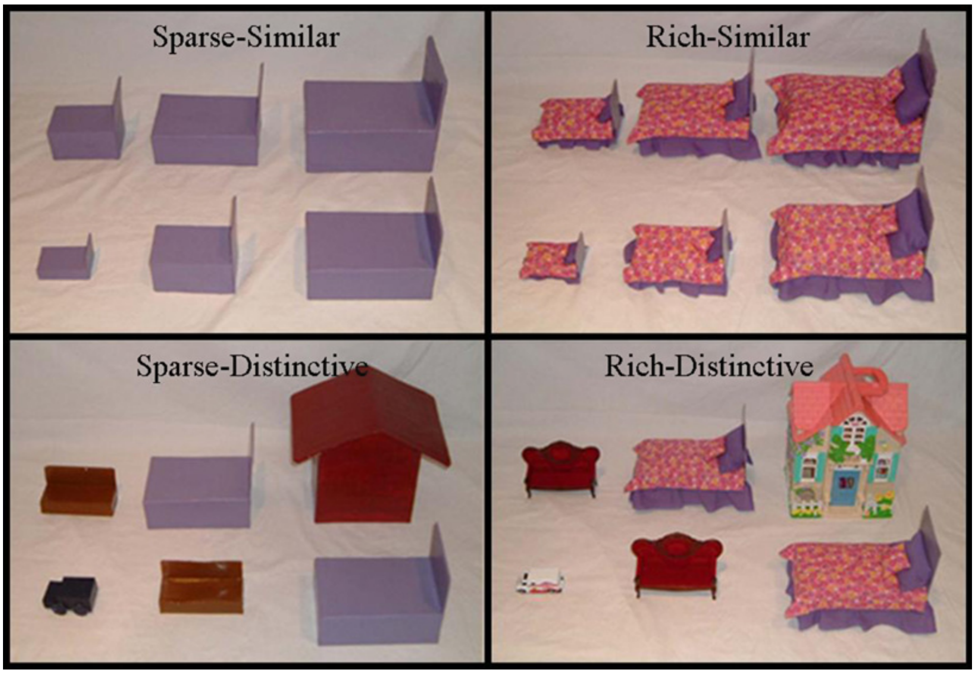

Fig. 4. Sample stimuli conditions in Experiment 2.

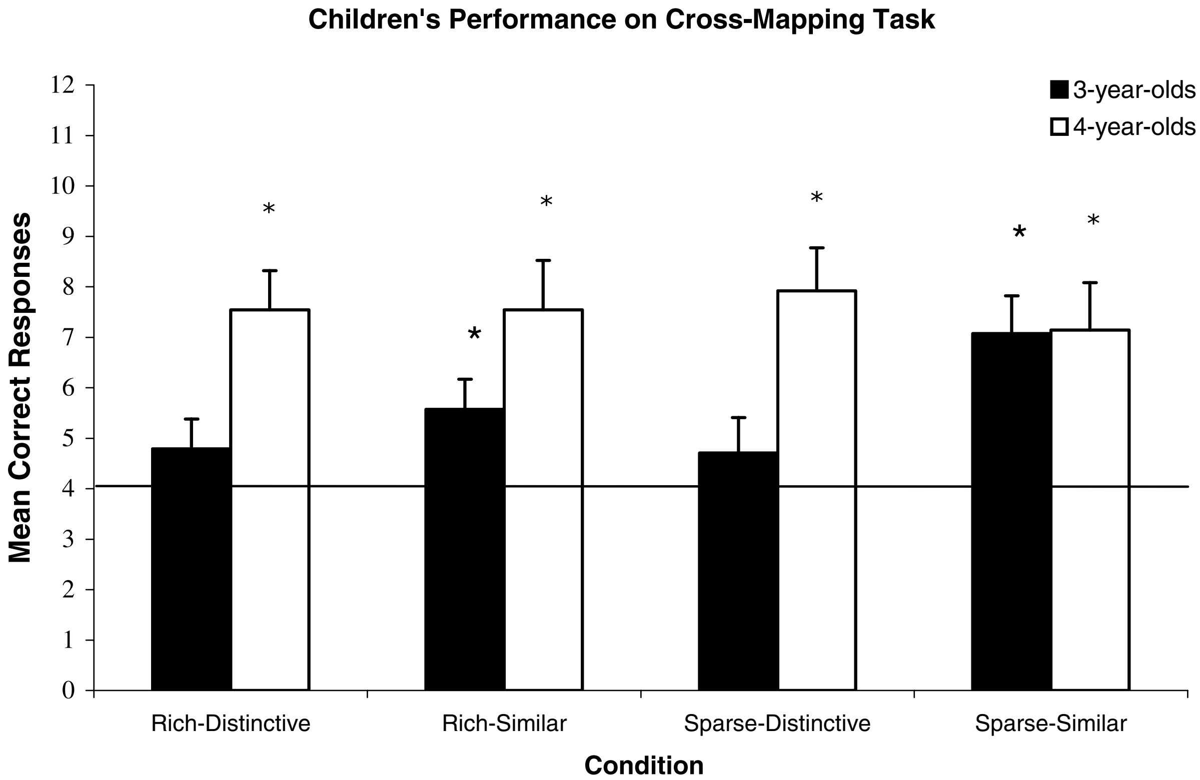

Fig. 5. Children's mean correct responses for the different stimuli conditions in Experiment 2. *Above chance, *ps* < .08.

were significantly higher than within-match ratings, $t(17) = 2.37$, $p < .03$, because one-third of the foils were identity matches (hence the cross-mapping of object identity with relative size). Interestingly, the similarity ratings did not differ for rich objects versus sparse objects (rich objects: $M = 6.60$, $SD = 3.22$; sparse objects: $M = 6.67$, $SD = 3.04$), $t(35) = 1.13$. The ratings were also comparable for the two object sets used in this experiment (Object Set 1: $M = 7.00$, $SD = 2.95$; Object Set 2: $M = 6.20$, $SD = 3.26$), $t(35) = 1.13$.

Children's mapping performance

All children completed the familiarization trials. The mean number of correct responses for each age group and condition are presented in Fig. 5. Unlike the task used in Experiment 1, the cross-mapping comparison pitted object similarity against relational similarity. Thus, children's performance should be much less accurate, particularly if they pay undo attention to object similarity. Consistent with this prediction, children's scores on the cross-mapping task were lower overall ($M = 5.53$, $SD = 2.60$) than they were on the identity match used in Experiment 1 ($M = 7.44$, $SD = 3.15$), $t(206) = 4.26$, $p < .001$. Also, as demonstrated previously by Gentner and Rattermann (1991), there was a significant difference between 3-year-olds' performance in the rich–distinctive condition ($M = 4.79$, $SD = 2.22$) and their performance in the sparse–similar condition ($M = 7.07$, $SD = 2.38$), $t(13) = 2.38$, $p < .03$. As before, when the displays were detailed and distinctive, performance was less accurate, presumably because this encouraged responding based on surface features. However, our results suggest that this difference is due to differences in distinctiveness rather than richness per se. Specifically, whereas 3-year-olds' scores exceeded chance in both conditions where distinctiveness was low, they were at chance in the two conditions where distinctiveness was high. These effects did not appear to be mediated by richness.

This conclusion was confirmed by a three-way ANOVA with age (3-year-olds vs. 4-year-olds), distinctiveness (distinctive vs. similar), and richness (rich vs. sparse) as between-participants factors. Neither the main effect nor any of the interactions involving richness was significant, $Fs \leqslant 1.01$. There was a significant main effect of age, $F(1, 113) = 12.06$, $p < .01$, indicating that, as before, 4-year-olds performed better than 3-year-olds overall. However, the differences due to distinctiveness did not reach significance as they did in Experiment 1, $F(1, 113) = 1.81$, $p < .18$. Perhaps the cross-mapping task was so difficult, even under the most supportive conditions, that random guessing swamped these differences. As in Experiment 1, an ANOVA using arcsin transformed scores yielded the same pattern.

Taken together, the results of Experiments 1 and 2 indicate that high distinctiveness increased the perceived similarity between identical objects. When the task requires object matches, distinctiveness improves performance by highlighting the identical objects. And when the task requires children to ignore object appearance, distinctiveness worsens performance. Only children with a clear understanding of the shared relation can detect it even when distinctiveness is high. Younger children, for whom relational understanding is weaker, perform well when distinctiveness is low but cannot resist the pull to the object matches when distinctiveness is high.

Experiment 3

Our results so far indicate that the degree to which matching objects pop out (or differ) from other objects in the scene appears to dominate children's responses in the relational mapping task. Still, we would not claim that within-match similarity is unimportant. Instead, we hypothesize that children's perception of similarity is guided by the ratio of these two aspects rather than either aspect alone.

Unfortunately, the possible contributions of within-match similarity were difficult to assess in Experiments 1 and 2 because the object matches were uniformly high in similarity (i.e., identity matches). In Experiment 3, we wanted to see what would happen when within-match similarity was not at ceiling, that is, when both within-match and nonmatch similarity were low. By including such conditions, we can evaluate two specific hypotheses.

First, we can compare performance for varying degrees of within-match similarity where distinctiveness is held constant. More specifically, we can compare scores where both within-match and nonmatch similarity are kept low (hereafter referred to as the low–low condition) with scores where within-match similarity is high and nonmatch similarity is low (hereafter referred to as the high–low condition). A significant difference between these conditions would demonstrate that within-match similarity does exert an influence so long as distinctiveness (i.e., nonmatch similarity) does not change enough to equate the ratios.

Second, the ratio hypothesis makes a counterintuitive prediction, namely that differences in within-match similarity should *not* affect performance so long as nonmatch similarity also decreases. In other words, so long as the *ratio* remains the same, changes in within-match similarity should not matter. Thus, scores in the low–low condition of Experiment 3 should be the same as scores in the rich–distinctive condition of Experiment 1 where both within-match and nonmatch similarity were kept high (hereafter referred to as the high–high condition). This may seem unlikely because when two objects are less similar (i.e., in the low–low condition), it stands to reason that they would be matched less readily. But if we are correct that similarity is determined by context in combination with features of the objects themselves, then there may be situations where this is not the case.

Method

Participants

A total of 14 3-year-olds (mean age = 3 years 6 months, range = 3 years 0 months to 3 years 11 months) and 14 4-year-olds (mean age = 4 years 6 months, range = 4 years 0 months to 4 years 11 months) participated. An additional 2 children were excluded from the study because they refused to complete the task. Within each age group, there were 7 boys and 7 girls. These age groups were included to permit direct comparisons between the results of this experiment and those of Experiment 1. Children were recruited from a predominantly white middle-class population through newspaper advertisements and local birth announcements. All children spoke English as their primary language.

Materials and procedures

The procedure from Experiment 1 was followed here. Because we did not expect to find differences in richness (based on the results of Experiments 1 and 2), we used only rich objects to streamline the data collection process. The only other change was that both within-match and nonmatch similarity in this experiment were kept low. For example, the children's array might contain a small truck, a medium cup, and a large guitar, whereas the experimenter's array might contain a small shoe, a medium cell phone, and a large bed. (For a complete list of stimuli, see Fig. 6 and the Appendix A) If the experimenter hid a sticker under her medium object (e.g., the cup), then the children were to find a sticker under their medium object (e.g., the cell phone). Thus, although within-match similarity was much lower here than in Experiment 1, nonmatch similarity was also kelp low; in fact, it was the same as in the distinctive objects condition of Experiment 1. This means that the ratio of similarity in the current experiment is roughly equal to the ratio in the previous rich–similar condition, where both types of similarity were high.

Results

Adults' similarity ratings

As in the previous experiments, we confirmed that our similarity manipulations were objective by obtaining pairwise ratings from a group of adults. A group of 15 undergraduate and graduate students rated 18 pairs (i.e., 9 pairs within each of two different object sets) presented in a fixed random order. As expected, both the within-match and nonmatch similarity for objects used in this experiment were low and did not differ (within-match: $M = 2.35$, $SD = 0.38$; nonmatch: $M = 2.67$, $SD = 0.63$), $t(8) = 1.13$. The ratings also confirmed that both object sets obtained the same similarity ratings on average (Object Set 1: $M = 5.06$, $SD = 3.34$; Object Set 2: $M = 4.99$, $SD = 3.33$), $t(8) = 0.62$.

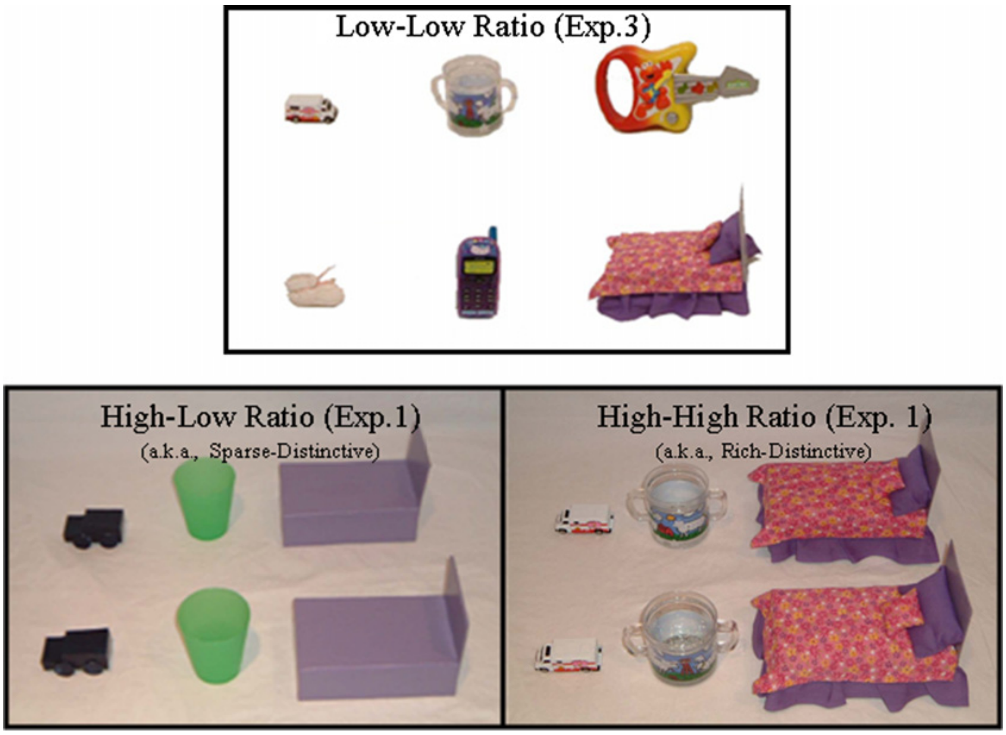

Fig. 6. Sample stimuli condition in Experiment 3.

In Experiment 1, both within-match and nonmatch similarity were high for the rich–similar condition (within-match: $M = 9.71$; nonmatch: $M = 8.40$). In contrast, both within-match and nonmatch similarity were low in the current experiment. Thus, although actual values of within-match and nonmatch similarity varied quite a lot between these two conditions, the *ratio* between within-match and nonmatch similarity was roughly 1:1 in both cases (i.e., 2.35/2.67 for the low–low ratio condition in Experiment 3 and 9.71/8.40 for the high–high ratio condition in Experiment 1). (For the ratios of conditions compared in Experiment 3, see Table 1.)

Children's mapping performance

All children completed the familiarization trials. Fig. 7 provides a comparison of children's mean scores in the low–low ratio condition of the current experiment with their mean scores in both the high–high ratio (i.e., rich–similar) and high–low ratio (i.e., rich–distinctive) conditions of Experiment 1. First, we compared performance between conditions with different levels of within-match similarity but with roughly equal nonmatch similarity.

Table 1
Similarity ratios of three conditions compared in Experiment 3

Condition	Within-match similarity rating	Nonmatch similarity rating	Ratio
High–high ratio (i.e., rich–distinctive, Experiment 1)	9.71	8.40	1.16
High–low ratio (i.e., rich–similar, Experiment 1)	9.70	2.35	4.13
Low–low ratio (Experiment 3)	2.35	2.67	0.88

Note. Ratios are based on adults' similarity ratings of all the stimuli pairs.

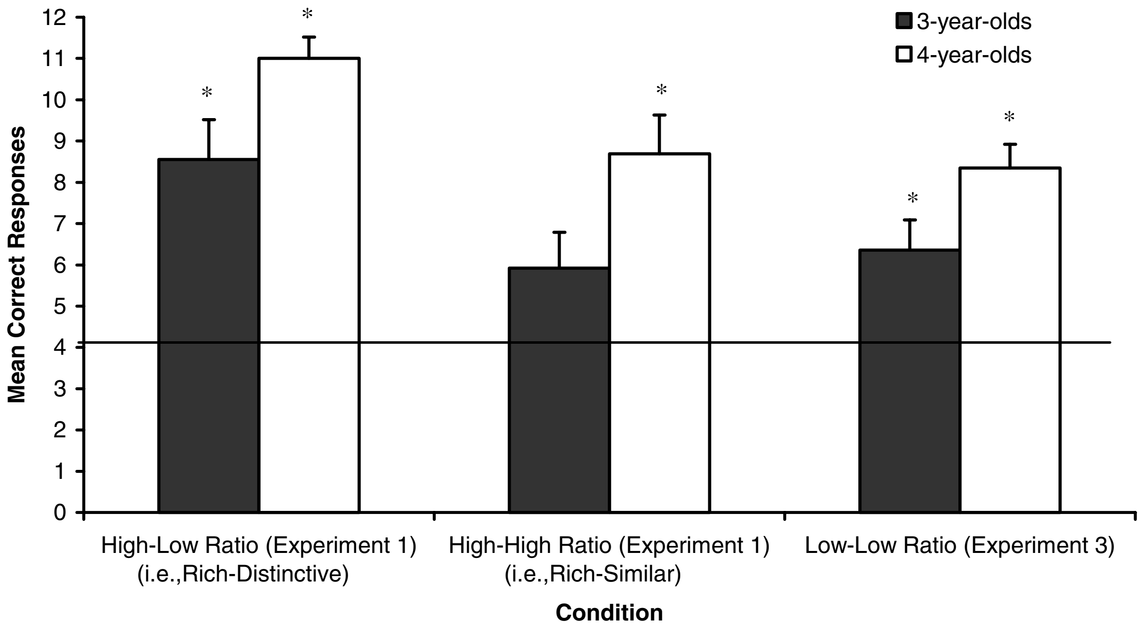

Fig. 7. Children's mean correct responses for the stimuli conditions in Experiments 3 and 1. *Above chance, $ps < .08$.

A two-way ANOVA with age (3- vs. 4-year-olds) and condition (high–low ratio vs. low–low ratio) as between-participants factors revealed that high similarity between matching objects does mediate comparisons when distinctiveness (i.e., nonmatch similarity) is held constant (high–low ratio: $M = 9.83$, $SD = 2.81$; low–low ratio: $M = 7.35$, $SD = 3.06$), $F(1, 76) = 5.58$, $p < .01$. There also was a significant effect of age where 4-year-olds performed better than 3-year-olds, $F(1, 76) = 12.80$, $p < .001$, but the interaction between age and condition was not significant.

Next, we compared performance between conditions with different levels of within-match and nonmatch similarity but roughly equal ratios. An ANOVA with age (3- vs. 4-year-olds) and condition (low–low ratio vs. high–high ratio) as between-participants factors confirmed that scores improve with age, $F(1, 56) = 8.01$, $p < .008$, but revealed no effect of condition, $F(1, 56) = 0.004$. Furthermore, no interaction between age and condition was found, $F(1, 56) = 0.78$, although both 3- and 4-year-olds performed above chance in the low–low condition (3-year-olds: $t(13) = 3.23$, $p < .008$; 4-year-olds: $t(13) = 5.21$, $p < .001$), whereas only 4-year-old performed above chance in the high–high condition. Thus, results suggest that children in both age groups essentially performed the same whether the within-match and nonmatch similarity was high or low. This is strong evidence that children's perception of similarity is determined by multiple factors. Even preschoolers who have demonstrated a strong attraction to high-similarity matches in past research do so only when the whole context points them in that direction.

General discussion

It has been hypothesized that children discover new dimensions of similarity in the process of making comparisons (Gentner & Medina, 1998; Gentner et al., 1995; Smith, 1989, 1993). If so, then an important question is what conditions lead children into these comparisons in the first place. In other words, what factors start this process in motion?

Previously, researchers focused on the effects of featural overlap between matching objects on children's comparisons (DeLoache, 1989, 1990; Gentner & Rattermann, 1991; Gentner & Toupin, 1986; Mandler & Bauer, 1988; Smith, 1984). However, the current results indicate that there is another important factor, namely the degree to which matching objects are distinct from other objects in the scene. The less the target objects overlap with the alternatives, the more likely children are to compare them. Thus, lack of overlap with distracter objects serves to increase the perceived similarity between the targets. In fact, our findings indicate that distinctiveness is so integral to children's similarity judgments that even extreme variations in within-match similarity do not affect performance unless the overall ratio permits it.

It is easy to imagine how changing similarity ratios could affect children's learning in naturalistic play. If a child took a red car, a red truck, a soccer ball, and a teddy bear out of the toy chest, it would set the stage for learning about "same color," "vehicle," or a range of other features shared between the truck and car. This would happen because, in the context of the other relatively dissimilar items (i.e., the soccer ball and the teddy bear), the overlap between the car and the truck would become more salient. If instead the child had pulled out a red car and three red trucks, comparisons among the trucks might be facilitated because in that case the trucks and the car have less overlap than the trucks have with each other. In this situation, the child may learn about truck subcategories or other commonalities among the trucks, such as size, that were not shared with the car. Importantly, these differences in learning would not be due to differences in the similarity between the car and the truck because that relation is not changing. Instead, it is the changing context that determines whether children will compare this pair of items or some other pairing in the scene. Thus, like high overall similarity, the contexts in which objects appear could be a critical factor in determining whether children will compare various items and, in turn, what they will take away from this experience.

Although our study is one of the first to demonstrate such effects in children's object comparisons, the idea that distinctiveness matters is not original to this work. Researchers have long recognized that perceived similarity depends on the context and frame of reference, a premise that has played a major role in formal theories of similarity and choice (Goldstone, Medin, & Gentner, 1991; Goldstone et al., 1997; Luce, 1959; Nosofsky, 1984; Tversky, 1977). In fact, when Markman and Gentner (1993) simulated a cross-mapping comparison using the structure-mapping engine, they found that the model generated a preference for the object matches based on the distinctiveness of the objects. Thus, the model predicts what we have found in the current study, namely that relational matches are more compelling when the matching objects are presented in contrasting contexts.

But how exactly does context exert this influence? By what mechanisms would the perception of similarity vary across contexts? Several theorists have suggested that shifts in perceived similarity reflect shifts in attention (Nosofsky, 1984; Shepard, 1962; Smith, 1989; Tversky, 1977). In any comparison, it is possible to focus on a range of different attributes and relations. Perhaps context constrains this focus by increasing attention to certain dimensions while decreasing attention to others. To illustrate, consider Smith's (1989) example of comparisons involving a bat, a crow, and a flamingo. If attention were given to shape and color, then the crow and the bat would move closer together in similarity space. However, if attention were shifted to the head and feet, then the crow and the flamingo would be closer in similarity space and the bat would move away. These shifts in attention would occur, according to Smith, in response to changing contextual information such as the presence of certain linguistic cues, changing task demands, and perceptual learning history. Indeed, this has been borne out (e.g., Deák et al., 2004, 2002; Jones et al., 1991; Kemler Nelson et al., 2000). On this view, even irrelevant features, such as glittery feet, might be enough to capture one's attention, thereby increasing the perceived similarity of the crow and flamingo.

Our results suggest that the features of the surrounding objects may also precipitate these attentional shifts. Just as in Tversky's (1977) study, where adults apparently shifted their attention from political to geographic features depending on the other choices presented in the comparison, children in our study appeared to shift their attention from some features to others based on distinctiveness. When the targets were highly distinctive from the alternatives, as in the comparison involving two medium beds alongside a small truck and a large house, attention to the shared "bedness" features, such as rectangular shape and fluffiness, may have increased. In contrast, when the targets were highly similar to their alternatives, as in the comparison involving only different-sized beds, attention to "bedness" would not increase and may even decrease because all of the objects in the scene shared these features.

This interpretation raises new questions about the flexibility of children's comparisons and the extent to which this differs from the flexibility exhibited by adults. As mentioned earlier, children's initial reliance on overall similarity could imply that their comparisons are rather rigid. However, the context effects obtained here and in previous research suggest that children's comparisons are not qualitatively different from those of adults. Like adults, children appear to shift their perception of similarity in response to changing situations. What may be changing developmentally, then, is the degree of sensitivity to context. Perhaps it takes more extreme contrasts to shift children's attention than it does to shift adults' attention. Alternatively, children's focus of attention may be more diffuse than that of adults, so the crucial developmental change involves a gradual narrowing of focus as more dimensions are learned. A third possibility is that early attention weighting may be driven by perceptual salience, whereas later in development controlled attention based on prior learning may be possible (Smith, 1989; Smith & Heise, 1992). Although the current results do not indicate which of these differences distinguish children from adults, they suggest that there is one important commonality, namely that the basic mechanism of shifting attention weights in response to changing context appears to be fundamentally the same.

The current results also raise questions about the role of overall similarity, and richness in particular, for children's comparisons. Gentner and Rattermann (1991) claimed that children perceive similarity more readily when matching items are richly detailed (i.e., have many points of alignment). We failed to obtain evidence of these differences when distinctiveness was controlled. On reflection, it is unclear how one could isolate richness in a complex display because even if rich items have more within-match similarity than the sparse items, they probably also have more nonmatch similarity. This means that richness differences may be obtainable in pairwise comparisons but would not necessarily affect behavior where ratio information is available.

In any case, it may be useful to revisit other research claiming to demonstrate richness effects in preschoolers' comparisons. In light of the current findings, it seems possible that such effects could have been mediated by the ratio of within-match similarity and nonmatch similarity rather than reflecting variations in within-match similarity alone. For example, in one version of DeLoache's (1989, 1990) classic model room experiments, 3-year-olds were significantly better at mapping a model onto a physical space when there was a great deal of surface similarity between the analogous items. In these experiments, children first watched as a miniature toy was hidden in a small-scale model room, and then they were asked to find a toy in the same location in a full-size room. When DeLoache varied the furniture items, children performed better when the items in the two rooms were rich (i.e., shared many surface features). However, the current results suggest that it may be important to vary the similarity of items within the rooms as well. Even with sparse objects, children may be able to see the underlying relations if the items in the rooms are highly distinctive from one another in shape, color, style, and so forth (e.g., a bed, a lamp, and a plant). The task may be more difficult if the items within each room are more similar (e.g., a bed, a desk, and a sofa) regardless of how detailed or rich the items are.

We conclude that children's comparisons are influenced by more than shared surface similarity. Instead, distinctiveness plays an important role. Further research that specifies the way different contexts affect children's perception of similarity is needed to fully understand the mechanism behind children's comparisons. However, the current findings and consistent context effects found in adults' similarity judgments indicate that this direction seems worth pursuing.

Acknowledgments

This research was supported in part by a Kemp grant and a National Institutes of Health developmental training grant (HD07475-07) to the first author and by a National Institutes of Health grant (HD37819-03) and a Spencer Foundation grant to the second author. We thank Brian Bowdle, Rob Goldstone, Susan Jones, and Linda Smith for helpful comments. Special thanks go to Jessica Kaufman, Alton Dorian Clark, and Zoe Williamson for their assistance with data collection and coding. Finally, we are most grateful to the participating children, parents, and schools without whose generous support and cooperation these experiments would not have been possible.

Appendix. Stimuli used in Experiments 1, 2, and 3

Experiment 1

Distinctive objects conditions (i.e., rich–distinctive and sparse–distinctive)

Set 1: 2 identical small trucks	Set 2: 2 identical small airplanes
2 identical medium sofas	2 identical medium mugs
2 identical large beds	2 identical large cars

Similar objects conditions (i.e., rich–similar and sparse–similar)

Set 1: 2 identical small beds	Set 2: 2 identical small cushions
2 identical medium beds	2 identical medium cushions
2 identical large beds	2 identical large cushions

Experiment 2

Distinctive objects conditions (i.e., rich–distinctive and sparse–distinctive)

Set 1: 1 small truck	Set 2: 1 small airplane
2 identical medium sofas	2 identical medium cups
2 identical large beds	2 identical large cars
1 extra large house	1 extra large cushion

Appendix (*continued*)

Similar objects conditions (i.e., rich–similar and sparse–similar)

Set 1: 1 small bed	Set 2: 1 small airplane
2 identical medium beds	2 identical medium cups
2 identical large beds	2 identical large cars
1 extra large house	1 extra large cushion

Experiment 3

Low–low ratio condition

Set 1: 1 small truck	Set 2: 1 small airplane
1 small shoe	1 small baby carriage
1 medium cup	1 medium bathtub
1 medium toy cell phone	1 medium pillow
1 large guitar	1 large book
1 large bed	1 large car

References

Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & Cognition, 10*, 82–93.

Deák, G. O., Ray, S. D., & Pick, A. D. (2004). Matching and naming objects by shape or function: age and context effects in preschool children. *Developmental Psychology, 38*, 503–518.

DeLoache, J. S. (1989). Young children's understanding of the correspondence between a scale model and a larger space. *Cognitive Development, 4*, 121–139.

DeLoache, J. S. (1990). Young children's understanding of models. In R. Fivush & J. Hudson (Eds.), *What children remember and why* (pp. 94–126). Cambridge, UK: Cambridge University Press.

Gati, I., & Tversky, A. (1984). Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology, 16*, 341–370.

Gentner, D., & Gunn (2001). Structural alignment facilitates the noticing differences. *Memory & Cognition, 29*, 565–577.

Gentner, D., & Markman, A. B. (1994). Structural alignment in comparison: no difference without similarity. *Psychological Science, 5*, 152–158.

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. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on thought and language: Interrelations in development* (pp. 225–277). Cambridge, UK: Cambridge University Press.

Gentner, D., Rattermann, M. J., Markman, A. B., & Kotovsky, L. (1995). Two forces in the development of relational similarity. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 263–313). Hillsdale, NJ: Lawrence Erlbaum.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277–300.

Goldstone, R. L., Medin, D. L., & Gentner, D. (1991). Relational similarity and the nonindependence of features in similarity judgments. *Cognitive Psychology, 23*, 222–262.

Goldstone, R. L., Medin, D. L., & Halberstadt, J. (1997). Similarity in context. *Memory & Cognition, 25*, 237–255.

Jones, S., Smith, L., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development, 62*, 499–516.

Kemler Nelson, D. G., Frankenfield, A., Morris, C., & Blair, E. (2000). Young children's use of functional information to categorize artifacts: three factors that matter. *Cognition, 77*, 133–168.

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

Luce, R. D. (1959). *Individual choice behavior*. New York: John Wiley.

Mandler, J. M., & Bauer, P. J. (1988). The cradle of categorization: Is the basic level basic? *Cognitive Development, 3*, 237–264.

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

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

Paik, J. H., & Mix, K. S. (2006). *It's all relative: different levels of relational similarity used in children's comparisons.* San Francisco State University. Submitted for publication.

Rips, L. J., & Collins, A. (1993). Categories and resemblance. *Journal of Experimental Psychology: General, 122*, 468–486.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology, 7*, 532–547.

Shepard, R. N. (1962). The analysis of proximities: multidimensional scaling with an unknown distance function (part 1). *Psychometrika, 27*, 125–140.

Sjöberg, L. (1972). A cognitive theory of similarity. *Göteborg Psychological Reports, 2*, 1–23.

Smith, L. B. (1984). Young children's understanding of attributes and dimensions: a comparison of conceptual and linguistic measures. *Child Development, 55*, 363–380.

Smith, L. B. (1989). From global similarity to kinds of similarity: the construction of dimensions in development. In S. Vosniadou & A. Orthony (Eds.), *Similarity and analogical reasoning* (pp. 146–187). Cambridge, UK: Cambridge University Press.

Smith, L. B. (1993). The concept of same. *Advances in Child Development and Behavior, 24*, 215–252.

Smith, L. B., & Heise, D. (1992). Perceptual similarity and conceptual structure. In B. Burns (Ed.), *Percepts, concepts, and categories: The representation and processing of information* (Advances in Psychology, No. 93, pp. 233–272). New York: Elsevier.

Soja, N. (1992). Inferences about the meanings of nouns: the relationship between perception and syntax. *Cognitive Development, 7*, 29–46.

Soja, N., Carey, S., & Spelke, E. (1991). Ontological categories guide young children's inductions of word meanings: object terms and substance terms. *Cognition, 38*, 179–211.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327–352.