Conceptual and Linguistic Biases in Children's Word Learning

Gil Diesendruck, Susan A. Gelman, and Kim Lebowitz
University of Michigan

Four studies examined the influence of essentialist information and perceptual similarity on preschoolers' interpretations of labels. In Study 1, 3-, 4-, and 5-year-olds were less likely to interpret 2 labels for animals as referring to mutually exclusive categories: when the animals were said to share internal, rather than superficial, properties and when the animals were perceptually similar rather than dissimilar. In Study 2, neither internal nor functional property information influenced 4-year-olds' interpretations of labels for artifacts. Studies 3 and 4 provide baseline data, demonstrating that the domain differences were not due to prior differences in children's lexical knowledge in the 2 domains. These results suggest that children have essentialist beliefs about animals, but not about artifacts, and that these beliefs interact with children's assumptions about word meaning in determining their interpretations of labels.

There has been much interest lately in the interaction among language, perception, and conceptual knowledge in children's language acquisition and conceptual development. A central question concerns how preschoolers decide that two entities are "of the same kind" and how this conceptual decision relates to children's naming decisions. Currently, researchers debate about whether children's labels are manifestations of deep conceptual structures (Soja, Carey, & Spelke, 1992) or whether they simply reflect children's bias to attend to perceptual properties of object (Barsalou, 1993; Smith, Jones, & Landau, 1996).

The purpose of the present studies is to address this broad question by investigating whether children rely on one kind of conceptual belief—essentialism—in extending labels. Also examined is whether this belief applies differently in various ontological domains and for varying degrees of perceptual similarity between entities in these domains.

Gelman, Medin, and their colleagues (Gelman, Coley, & Gottfried, 1994; Gelman & Diesendruck, in press; Medin & Ortony, 1989) claimed that people have a belief that certain categories, especially natural kinds, have some essential nonobvious properties that determine what a category is. Two lines of research provide evidence that children have essentialist beliefs about natural kinds. First, studies show that children believe the identity of an animal changes only if its internal, and thus presumably essential, properties are changed. Keil (1989), for instance, told kindergartners, second and fourth graders a story about a raccoon who underwent a transformation that made it look like a skunk (e.g., had its fur shaved and dyed). He found that by the second grade, children insisted that the animal was still a raccoon. Gelman and Wellman (1991) presented 4- and 5-year-olds with three kinds of transformations: (a) removal of insides (e.g., the stuff inside an animal, such as bones and blood), (b) removal of outsides (e.g., an animal's fur), and (c) movement. Children were then asked whether the item would still be the same one and whether it would still function after the transformation. As predicted, most children said that the identity and functioning of an animal would change if its insides were removed but not if its outsides were removed or if it moved. In sum, children seemed to assume that if animals have different essences, they must belong to different categories.

A second, more indirect source of evidence for essentialist beliefs in children stems from induction studies. Gelman and Markman (1986) showed children triads of animals and asked them to infer which two animals shared certain properties. They found that, a priori, children inferred that similar looking animals had common properties, even essential ones such as internal organs. However, when told that two perceptually dissimilar animals were members of the same category (i.e., had the same name), children inferred that these two animals had common internal properties, even though they did not look alike. In other words, children reasoned that if animals had the same name, they must have the same internal properties.

However, one crucial implication of essentialist reasoning in children has yet to be explored. In fact, it is perhaps the most direct implication of essentialism. Specifically, children should infer that if two entities have the same essence, then they are of the same kind and thus have a common name. In other words,
we should expect children to extend labels to objects primarily on the basis of whether they believe that the objects have the same essence. One of the reasons this has not been tested previously is partly because children may not know specific essences but instead have "essence placeholders" (Medin & Ortony, 1989). However, as noted earlier, internal parts may be a good proxy for essences. At the very least, they appear to be relatively more essential on tasks of induction and identity judgment.

Any test of conceptual influences on children's naming decisions must take into account children's word learning principles or biases (Golinkoff, Mervis, & Hirsh-Pasek, 1994; Markman, 1989; Merriman & Bowman, 1989). Thus, any account of how conceptual information influences children's naming is, in fact, an interactionist account; that is, an explanation of how such sources of information modulate the operation of lexical biases. One lexical bias is particularly pertinent in the present case because of its implications for decisions about category membership—Markman's mutual exclusivity bias. According to this bias, children assume that an object can have only one label—that, thus, different labels pick out mutually exclusive categories.

In the present studies, we capitalize on this lexical assumption and the labeling decisions caused by it and use it to provide a test of essentialist reasoning in children. The basic logic of our test is the following: We give children two kinds of information about different entities. First, we convey to children whether the entities have the same internal parts (as a proxy for essences); then, we give different labels for the entities. As described above, essentialism should lead children to believe that entities with the same essential properties also share a name. Mutual exclusivity, however, should lead children toward the opposite conclusion: that entities with different names should be of different kinds. Children's responses in this context should tell us whether and how essentialist beliefs modulate the operation of the mutual exclusivity bias.

Operationally, we had to label the entities in such a way that it would at least be plausible to extend one label to different entities. For this purpose, we used a labeling procedure similar to the one used by Gelman, Wilcox, and Clark (1989). For instance, we taught children a novel label for an atypical exemplar of a familiar category (e.g., pointing to a flying squirrel: "This one is a squirrel; it's a mef.") and then labeled a typical exemplar of the familiar category (e.g., pointing to a typical squirrel: "This one is a squirrel; it's not a mef."). Even though this labeling input implies an overlap in the extension of the novel and familiar labels, in Gelman et al., preschoolers overwhelmingly interpreted the two labels as picking out mutually exclusive categories. In other words, children extended the novel label only to the atypical exemplar but erroneously narrowed the extension of the familiar label and applied it only to the typical exemplar (see also Merriman & Stevenson, 1997). As Merriman and Bowman (1989) noted, the mutual exclusivity bias affects not only the interpretation of the meaning of a novel label but also the relation between a novel label and other labels in the child's lexicon. Specifically, as was the case in Gelman et al., mutual exclusivity may lead children to correct or restrict their extension of a familiar label to preserve nonoverlapping extensions (what Merriman & Bowman referred to as the "correction" and "restriction" effects of mutual exclusivity). The present studies test whether essentialist beliefs help children override mutual exclusivity by leading them not to correct or restrict the extension of a familiar label after having learned a novel label.

Study 1 tests whether children's essentialist beliefs about animals underlie their naming decisions. Half of the children were told that two animals share internal properties (thus presumably essential), and the other half were told that the animals share superficial properties. The pairs of animals presented to children varied in their perceptual similarity. The experimenter then labeled the animals, as described above. We hypothesized that children receiving internal property information would be less likely to interpret the novel and the familiar labels as mutually exclusive compared with children receiving superficial information. We further expected children to rely on the perceptual similarity between animals as a heuristic for category membership judgments and thus apply a mutual exclusivity interpretation more often for dissimilar animals than for similar animals. It is important, nonetheless, that we expected the effect of property information to be robust and to hold for both similar and dissimilar pairs of animals.

In Study 2, we extended this investigation to the domain of artifacts to test the domain-generality of children's essentialist beliefs. For that study, in addition to superficial and internal property information, we provided children with functional property information in which they were told about common uses of two objects. Some researchers have noted the centrality of function in the definition of artifacts (e.g., Keil, 1986; Nelson, 1995). It was tenable, then, that children would treat functional properties of artifacts as essential to determining an object's label. Study 3 served as a control by examining children's word learning when no property information about the animals or artifacts was provided. Study 4 examined children's initial extensions of the names of the animals and artifacts presented.

In terms of the mutual exclusivity bias per se, the present studies presuppose, as Woodward and Markman (1991) claimed, that the bias may work as a default assumption that can be overridden when children are convinced that the referents of labels are of the same kind. For instance, studies have found that 2-year-olds are less likely to treat two labels as mutually exclusive when they are sure that the referents are perceptually similar than when they are dissimilar, a pattern that could be interpreted as a relaxation of the mutual exclusivity bias (Diesendruck & Slatz, 1998; Waxman & Senghas, 1992). The present studies analogously test whether essentialist beliefs can lead children to conceive of two entities as belonging to the same category and thus help them overcome a mutual exclusivity bias. The hypothesis, then, is that children should maintain that the entities belong to mutually exclusive categories in the following conditions: (a) when told that two animals share nonessential properties (as in the superficial condition of Study 1), (b) when the entities being labeled are not believed to have essences altogether (as may be the case for artifacts—Study 2), or (c) when not provided with any property information (Study 3). (On this last point, see Waxman, 1990; Waxman, Shipley, & Shepperson, 1991, for evidence that presenting novel labels alone, in contrast to presenting labels in conjunction with prop-
Study 1

Method

Participants

Seventy-nine children participated in this study: 21 three-year-olds (M = 3 years 8 months, range = 3 years 3 months to 3 years 11 months), 32 four-year-olds (M = 4 years 7 months, range = 4 years 0 months to 4 years 11 months), and 26 five-year-olds (M = 5 years 7 months, range = 5 years 1 month to 6 years 1 month). There were 41 boys and 38 girls, distributed in approximately equal proportions across the three groups. Two additional children (one 3-year-old and one 4-year-old) were dropped from the study for failing to complete the task. Half of the 4- and 5-year-olds participated in each of the two conditions; 10 three-year-olds participated in the superficial condition, and 11 three-year-olds participated in the internal condition. The mean age of children in the two conditions did not differ significantly. All participants attended one of two preschools affiliated with a midwestern university. They constituted a racially and ethnically mixed, middle-class sample. Ten undergraduate students provided ratings of the stimuli.

Design

The experiment was a 2 (condition: internal, superficial) x 3 (age: 3-, 4-, and 5-year-olds) x 2 (perceptual similarity: dissimilar, similar) x 2 (order: familiar first, novel first) design. Condition and age were between-subjects variables; perceptual similarity and order were within-subjects variables.

Stimuli

The stimuli for this experiment were color photographs or realistic color drawings of animals, approximately 10 x 10 cm in size. Each child saw 10 sets of five pictures each. The pictures were arranged in an album such that on the left-hand pages, there was one picture of each of the two target animals, and on the right-hand pages, there were the same two pictures of the target animals plus a picture of a “distractor” animal placed between them. The distractor was included to make sure children were not responding randomly. The position of the target pictures on the pages alternated. A picture that was on the top in the left-hand page was placed on the bottom in the right-hand page. A cardboard flap separated the pages of the album so that the pictures on the right and on the left were not visible at the same time. The album also included one warm-up set (a blue square, a red square, and a red triangle) and two filler sets (ball, apple, banana, bicycle, motorcycle, boat).

The physical similarity between the two target pictures within each set was systematically varied. Each participant received five similar sets and five dissimilar sets (see Figures 1A and 1B for examples). The perceptual similarity of the pictures was established by adult ratings. Adult participants were instructed to judge the degree to which the two animals looked alike and were asked to take into consideration features such as shape, size, and complexity. They recorded their answers on a 1-to-7 scale, 1 being not at all similar and 7 being extremely similar. The mean score for dissimilars was 2.72; the mean for similars was 5.98. This difference was significant, t(8) = 12.00, p < .0001 (see Table 1).

Procedure

The experimenter saw all children individually. Children were introduced to a puppet and were told that they were going to look at a book with pictures of animals. The experimenter explained that some of the names of the animals were words the puppet knew. This was done so that children would not think they were being taught actual English words.

The first page of the album was the training set. Children were asked “Is there a square?” and “Is there another square?” They were then asked “Is there a red thing?” and then “Is there another red thing?” Children were corrected when necessary. This set was included to convey to children that they could pick the same picture in response to two different questions and could pick two pictures in response to one question (both of which were necessary to perform correctly on the experimental items).

The first set of animal pictures followed the training set. The order of sets was randomized and was the same for all children (snake, squirrel, rhinoceros, lizard, mouse, giraffe, cat, whale, dog, frog). As described in the Stimuli section, left-hand pages contained the two pictures of the target animals: a typical and an atypical exemplar of each category. The experimenter first provided information about features that the two animals shared, according to the condition to which the child was assigned. In the internal condition, the experimenter told the children that “This one [pointing to the atypical animal] has the same stuff inside as this one [pointing to the typical animal]. It has the same kind of bones, blood, muscles, and brain that this other one has.” In the superficial condition, children were told that “This one [pointing to the atypical animal] is the same size as this one [pointing to the typical animal]. It lives in the same zoo in the same kind of cage as this other one does.”

The experimenter then labeled the two animals in counterbalanced order. The atypical exemplar was always labeled with the familiar name followed by the novel name. For example, “This one [pointing to the atypical squirrel] is a squirrel; it’s a mef.” The typical exemplar was always labeled with the familiar name and then was said not to have the novel name. For example, “This one [pointing to the typical squirrel] is a squirrel; it’s not a mef.” After naming the pictures, the experimenter opened the cardboard flap and revealed three pictures on the right-hand page: the same two the child had just seen and an unrelated animal. The experimenter then asked the children to identify the animals. For instance, “Can you show me a mef? Is there another mef?” This latter prompt was repeated until the child said, “No.” Children were then asked about the other label, in this example, “squirrel.”

We hypothesized that the order of questioning of the labels would pragmatically influence how children responded to the experimenter’s naming requests. Specifically, we believed that asking for the novel label first and for the familiar label second (as illustrated above) would reinforce children’s tendency to narrow their extension of the familiar label to maintain mutual exclusivity. After all, by the time they were asked to pick referents of the familiar label (e.g., “squirrel”), children would have already committed to one of the animals as being the referent of the novel label (e.g., the atypical squirrel for “mef”) and thus would be drawn to exclude that animal from the extension of the familiar label. Conversely, when asked for the familiar label first, children would not have any explicit pressure to correct or restrict the extension of the familiar label and thus would be “free” to pick both animals as its referents.

To test this pragmatic effect, we counterbalanced the order in which the animal labels were introduced and requested within subjects. For items 1, 3, 5, 7, and 9, all children were introduced first to the typical animal in the pair (e.g., “This is a snake; it’s not a zav.”) and then to the atypical (“This is a snake; it’s a zav.”). Subsequently, children were...
asked first to identify referents of the familiar label (e.g., "snake") and then referents of the novel label (e.g., "zav"). We refer to this as Order 1. For items 2, 4, 6, 8, and 10, children were introduced first to the atypical animal and afterward to the typical animal (e.g., "This is a squirrel; it's a mef." and "This is a squirrel; it's not a mef."). Subsequently, children were asked first to identify referents of the novel label (e.g., "mef") and then referents of the familiar label (e.g., "squirrel"). This was defined as Order 2. Following the rationale presented above, we predicted that children would make more mutual exclusivity interpretations in Order 2 than in Order 1.

Two filler sets were placed after the fourth and seventh experimental sets. They were introduced to present some novelty and to help maintain children's interest in the task. Children were simply asked to help the experimenter name the pictures in each set (e.g., "ball," "apple," "banana").

### Table 1

**Items Used in Study 1**

<table>
<thead>
<tr>
<th>Category</th>
<th>Artificial labels used</th>
<th>M perceptual similarity between typical and atypical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Similars</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Snake</td>
<td>Zav</td>
<td>6.2</td>
</tr>
<tr>
<td>Rhinoceros</td>
<td>Kiv</td>
<td>6.0</td>
</tr>
<tr>
<td>Mouse</td>
<td>Jop</td>
<td>6.1</td>
</tr>
<tr>
<td>Whale</td>
<td>Lorse</td>
<td>6.0</td>
</tr>
<tr>
<td>Frog</td>
<td>Vit</td>
<td>5.6</td>
</tr>
<tr>
<td>Dissimilars</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squirrel</td>
<td>Mef</td>
<td>3.0</td>
</tr>
<tr>
<td>Lizard</td>
<td>Sap</td>
<td>2.8</td>
</tr>
<tr>
<td>Giraffe</td>
<td>Zon</td>
<td>2.5</td>
</tr>
<tr>
<td>Cat</td>
<td>Wag</td>
<td>3.4</td>
</tr>
<tr>
<td>Dog</td>
<td>Cak</td>
<td>1.9</td>
</tr>
</tbody>
</table>

**Results**

Our main hypothesis was that internal property information would lead children to conceive of two animals as being of the same kind and thus having a common name. We further hypothesized that perceptual similarity and order would influence how children interpreted the labels of the animals. These hypotheses led to two separate but related sets of predictions. First, the factors could influence how children extended the familiar labels, the novel labels, or both kinds of labels. For instance, children in the internal condition could reject applying the novel label to any of the animal exemplars, reckoning that if two animals have common internal properties, they cannot have different names. On the basis of Gelman et al.'s (1989) findings, however, we expected the factors would influence primarily how children extended the familiar labels, independently of whether they learned the novel label. Second, and more important, we expected that having learned the novel label, condition, perceptual similarity, and order would influence how children interpreted the relation between the familiar and the novel labels. In general, we predicted that children would be more likely to extend the familiar label to both animals and would be less likely to interpret the relation between the labels as one of mutual exclusivity when (a) told the animals share internal properties, (b) the animals were similar, and (c) asked for the familiar label first.

**Coding**

Children's responses were coded regarding the type of semantic relation they maintained between the novel and the familiar labels. These interpretation patterns also varied in terms of
whether children correctly extended the familiar label (to both the typical and the atypical exemplars) and the novel label (to only the atypical exemplar). These responses are illustrated in Table 2.

**Subordinate.** Subordinate (a) responses included two types: a - 1 (correct both), in which the child selects both the typical and atypical exemplars as referents of the familiar label, and only the atypical exemplar as the referent of the novel label, and a - 2 (correct familiar), in which the child selects both the typical and atypical exemplars as referents of the familiar label and only the typical exemplar as the referent of the novel label.

**Mutual exclusivity.** Mutual exclusivity (b) responses included the following: b - 1 (correct novel), in which the child picks only the typical exemplar as the referent of the familiar label and only the atypical exemplar as the referent of the novel label, and b - 2 (incorrect both), in which the child picks only the atypical exemplar as the referent of the familiar label and only the typical exemplar as the referent of the novel label.

**Superordinate.** In superordinate (c; incorrect both) responses, the child chooses only the typical or the atypical exemplar as the referent of the familiar label and both exemplars as referents of the novel label.

**Synonymy.** Synonymy (d; correct familiar) refers to a response in which the child picks both exemplars in response to both labels.

**Other.** Other (e) refers to the following types of responses: e - 1 (correct familiar), in which the child selects both exemplars as referents of the familiar label but none or the distractor as referent of the novel label; e - 2 (correct novel), in which the child selects the atypical exemplar as referent of the novel label but none or the distractor as referent of the familiar label; and e - 3 (incorrect both), in which the child selects the typical exemplar as referent of the familiar label and the distractor as referent of the novel label.

Table 2 presents the mean number of responses of each type as a function of condition and age. The first set of analyses examined the extent to which children correctly extended the familiar and the novel labels. The second set of analyses examined the effects on the type of relation children established between the novel and familiar labels.

**Label Extensions**

Two measures were created for this first set of analyses. The first measure was the number of correct interpretations of the familiar label. This measure consisted of the sum of responses a - 1, a - 2, d, and e - 1. The second measure was the number of correct interpretations of the novel label. This measure consisted of the sum of responses a - 1, b - 1, and e - 2 (see Table 3). Separate repeated-measures analyses of variance (ANOVA) were conducted on the proportion of correct familiar and correct novel interpretations. In these analyses, condition and age were entered as between-subject variables and perceptual similarity and order as within-subjects variables. For these analyses, as for all other analyses in which the original dependent measures were proportions, the proportions were transformed with an arcsine function before being entered into the ANOVAs. Proportions were used instead of means because of the unequal number of trials in the cells of the $2 \times 2$ (Perceptual Similarity $\times$ Order) matrix.

As expected, condition had a significant effect on the number of correct interpretations of the familiar label, $F(1, 73) = 4.64, p < .05$. Children were more likely to correctly extend the familiar label to the two animals when they were asked for the novel label first (Order 1, $M = 65\%$) than when asked for the familiar label first (Order 2, $M = 40\%$), $F(1, 73) = 72.25, p < .0001$. The only significant interaction was between order and perceptual similarity, $F(1, 73) = 9.67, p < .005$. Children were more likely to extend the familiar label to both animals when the animals were similar ($M = 61\%$) than when they were dissimilar ($M = 45\%$). Also as predicted, children were more likely to extend the familiar label to both exemplars when asked for the familiar label first (Order 1, $M = 65\%$) than when asked for the novel label first (Order 2, $M = 40\%$), $F(1, 73) = 72.25, p < .0001$. The only significant interaction was between order and perceptual similarity. $F(1, 73) = 4.45, p < .05$. Age did not have a significant effect on this measure.

As for how well children interpreted the novel label, condition, age, and order did not have significant effects, but perceptual similarity did. Children were more likely to extend the novel label correctly when the animals were similar ($M = 76\%$) than when they were dissimilar ($M = 71\%$), $F(1, 73) = 4.75, p < .05$. The interaction between order and perceptual similarity was also significant, $F(1, 73) = 7.07, p < .05$. None of the other interactions were significant.

**Relation Between Labels**

For the second set of analyses, we focused primarily on the subordinate and mutual exclusivity response types (i.e., responses a and b described in the Coding section). A subordinate response was the semantic relation implied by the experimenter's labeling input. With such a response, the child expressed...
Table 3

<table>
<thead>
<tr>
<th>Response type</th>
<th>Study 1</th>
<th></th>
<th>Study 3</th>
<th></th>
<th>Study 2</th>
<th></th>
<th>Study 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intern</td>
<td>Superf</td>
<td>label</td>
<td>Intern</td>
<td>Superf</td>
<td>label</td>
<td></td>
</tr>
<tr>
<td>Subordinate (a - 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>3.1</td>
<td>1.64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-year-olds</td>
<td>4.62</td>
<td>2.94</td>
<td>1.6</td>
<td></td>
<td></td>
<td>2.69</td>
<td>2.38</td>
</tr>
<tr>
<td>5-year-olds</td>
<td>3.08</td>
<td>3.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subordinate (a - 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>0</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-year-olds</td>
<td>0.38</td>
<td>0.06</td>
<td>0.13</td>
<td></td>
<td></td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td>5-year-olds</td>
<td>0.31</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutual Exclusivity (b - 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>3.7</td>
<td>5.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4-year-olds</td>
<td>2.81</td>
<td>5.5</td>
<td>5.27</td>
<td></td>
<td></td>
<td>3.85</td>
<td>4.56</td>
</tr>
<tr>
<td>5-year-olds</td>
<td>3.54</td>
<td>3.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mutual Exclusivity (b - 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>0.3</td>
<td>0.36</td>
<td>0.6</td>
<td></td>
<td></td>
<td>0.46</td>
<td>0.44</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>0.12</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-year-olds</td>
<td>0.38</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Superordinate (c)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>0.3</td>
<td>0.27</td>
<td>0.2</td>
<td></td>
<td></td>
<td>0.15</td>
<td>0.25</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>0.12</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-year-olds</td>
<td>0.15</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synonymy (d)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>1.5</td>
<td>1.27</td>
<td>1.27</td>
<td></td>
<td></td>
<td>2.31</td>
<td>1.69</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>1.56</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-year-olds</td>
<td>2.23</td>
<td>0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (e - 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>0.8</td>
<td>1.09</td>
<td>0.67</td>
<td>0.15</td>
<td>0.5</td>
<td>0.33</td>
<td>0.44</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>0.31</td>
<td>0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-year-olds</td>
<td>0.15</td>
<td>0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (e - 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>0.1</td>
<td>0</td>
<td>0.23</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>0</td>
<td>0</td>
<td></td>
<td>0.15</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-year-olds</td>
<td>0.15</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (e - 3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-year-olds</td>
<td>0.2</td>
<td>0.27</td>
<td>0.27</td>
<td>0.08</td>
<td>0.12</td>
<td>0.2</td>
<td>0.06</td>
</tr>
<tr>
<td>4-year-olds</td>
<td>0.06</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5-year-olds</td>
<td>0</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Intern = internal condition; superf = superficial condition; funct = functional condition.

an understanding that both animals were members of the same basic-level category and that the novel label denoted a subtype of the basic-level category. On the basis of previous research, and from a theoretical standpoint, mutual exclusivity responses seemed the most likely and interesting alternative response. With such a response, children indicated that they were not convinced that the two animals belonged to the same category and thus, by default, assumed the labels denoted mutually exclusive categories.

For the statistical analyses, we summed the number of responses of each type, namely subordinate or mutual exclusivity, that children gave. Separate repeated-measures ANOVAs were then conducted with the arcsine transformed proportion of responses of a single type (subordinate or mutual exclusivity) out of all responses as the dependent measures. In these analyses, age and condition were included as between-subjects variables, and perceptual similarity and order as within-subjects variables.

Preliminary analyses revealed no effect of gender on these measures.

Condition did not have a significant effect on the amount of subordinate interpretations children made, $F(1, 73) = 1.74, p > .10$, although there was a tendency for children in the internal condition to make more subordinate responses than children in the superficial condition (see Table 3). Perceptual similarity also did not have a significant effect on how likely children were to interpret the novel label as subordinate to the familiar one, $F(1, 73) = 2.16, p > .10$, though the difference between similars ($M = 40\%$) and dissimilars ($M = 29\%$) was in the direction predicted (see Figure 2). Neither the effect of age nor the interactions among these variables reached significance.

Nonetheless, order had the expected effect on the proportion of subordinate interpretations. Children were more likely to make subordinate interpretations when they were questioned about the familiar label first (Order 1, $M = 43\%$) than when
they were questioned about the novel label first (Order 2, \( M = 26\% \)), \( F(1, 73) = 22.66, p < .0005 \) (see Figure 2). The three-way interaction among order, perceptual similarity, and condition was also significant, \( F(1, 73) = 4.16, p < .05 \). The two-way interactions between order and condition, order and age, and order and perceptual similarity were not significant.

As for mutual exclusivity responses, condition did have the expected effect. Children made more mutual exclusivity interpretations in the superficial than in the internal condition, \( F(1, 73) = 5.75, p < .02 \). Furthermore, children were more likely to interpret a novel label as referring to a mutually exclusive category when its referent was perceptually dissimilar from the typical exemplar (\( M = 51\% \)) than when it was similar (\( M = 38\% \)), \( F(1, 73) = 11.90, p < .001 \). Age did not have a significant effect, and the interactions between perceptual similarity and any of the between-subjects variables were not significant.

Again, we found an effect of order in the hypothesized direction. Children were more likely to make mutual exclusivity interpretations when they were asked for the novel label first (Order 2, \( M = 55\% \)) than when they were asked for the familiar label first (Order 1, \( M = 34\% \)), \( F(1, 73) = 55.35, p < .0001 \). The interaction between order and perceptual similarity was significant, \( F(1, 73) = 28.73, p < .0001 \). Individual \( t \) tests were conducted to analyze this interaction. To control for the number of follow-up \( t \) tests conducted, we multiplied the \( p \) values by the number of tests (4). As illustrated in Figure 2, there was a significant effect of order for similar items, \( t(78) = 8.50, p < .05 \), but not for dissimilar items, \( t(78) = 0.74, p > .05 \). The three-way interaction among order, perceptual similarity, and condition was significant, \( F(1, 73) = 5.43, p < .05 \), though uninterpretable. There were no two-way interactions between order and any of the between-subjects variables.

Analyses of variance were conducted on all the other response types (i.e., synonymy, superordinate, and other), with condition and age as between-subjects variables. Repeated-measures ANOVAs could not be used on these responses because of too many empty cells. The only significant effect in all these analyses was one of age on the amount of "other" responses. Younger children gave this type of response more often than older children, \( F(1, 73) = 6.79, p < .005 \).

**Individual Response Patterns**

Analyses of the effect of condition were also conducted on children's individual response patterns. For this purpose, children were classified according to whether they gave the same response type for more than half the items. This was chosen as a criterion to look for the effect of condition above and beyond the effects of perceptual similarity or order (i.e., half of the items were perceptually similar and the other half dissimilar, so we could expect a certain response type for exactly half of the items simply because of perceptual similarity alone). Of the 39 children in the internal condition, only 13 gave a mutual exclusivity response on more than half the 10 items. Of the 40 children in the superficial condition, 23 did so, \( \chi^2(1, N = 79) = 4.65, p < .05 \). For subordinate interpretations, the pattern was the reverse. Fourteen of the 39 children in the internal condition gave subordinate responses for more than half the items, whereas only 7 of the 40 children in the superficial condition did so, \( \chi^2(1, N = 79) = 3.42, p = .064 \). Overall, as predicted, children in the internal condition seemed to be more likely to accept that two animals shared the same label and were less likely to treat these labels as mutually exclusive than were children in the superficial condition.

As reported above, there were no main effects of age on the types of responses children provided. However, older children were more likely than younger children to maintain one response type for many of the items. To capture this developmental trend, we divided children into two groups according to whether they gave the same type of response, no matter which, for more than half of the 10 items. Of the 21 three-year-olds, 13 (62%) gave the same type of response for more than half of the 10 items compared with 28 of the 32 (87%) 4-year-olds and 24 of the 26 (92%) 5-year-olds who did so, \( \chi^2(2, N = 79) = 8.37, p < .05 \).

**Discussion**

The present study provides evidence that preschoolers consider conceptual information when deciding on the extension of
and the semantic relation between names of animals. Children who were told that two animals shared internal (and thus presumably essential) properties were more likely to extend a common label to both animals, and were less likely to interpret labels for these animals as picking out mutually exclusive categories, than children who were told that the animals shared superficial properties. The fact that age did not interact with the effect of property information supports Gelman et al.'s (1994) contention that from an early age, children conceive of internal properties of animals as essential to their identities.

Consistent with previous studies (e.g., Diesendruck & Shatz, 1998; Smith et al., 1996; Waxman & Senghas, 1992), we found that children were affected by the degree of perceptual similarity between the referents of labels. Children were more likely to extend a familiar label to two perceptually similar animals and were, furthermore, less likely to assume that two labels referred to mutually exclusive categories when the animals were perceptually similar than when they were dissimilar. This finding is particularly striking given that perceptual similarity was a within-subjects variable; in other words, throughout the 10-set task, children somewhat modulated their interpretation of the labels according to item similarity. There was no interaction between perceptual similarity and age. That is, younger children were not more strongly influenced by the perceptual similarity between animals than older children. More important, there was no interaction between perceptual similarity and condition. Thus, as predicted, information about internal properties drove children to infer that two animals shared a common label even when the animals did not look alike.

Children were also influenced by the order of testing of the labels, suggesting that, as hypothesized, they were influenced by pragmatic considerations when deciding on the extension of labels. An important point to emphasize, though, is that order did not interact with condition for any of the response types. That is, the effect of condition on children's responses was independent of the order of questioning of the labels.

An unanticipated finding was that older children were more consistent than younger children—specifically, they were more likely to maintain one response type, no matter which, for 6 or more of the 10 items. It is interesting, however, that age had no effect on how correct their interpretations were. It seems that with age, children become more likely to commit to a rule or to an interpretation strategy (see also Ravn & Gelman, 1984; Siegler, 1983).

Together, the findings from this study suggest that by age 3 years, children use multiple sources of information to decide on the semantic relation between labels. Language (i.e., lexical biases), perception (i.e., degree of perceptual similarity), conceptual knowledge (i.e., internal vs. superficial information), and pragmatic considerations (i.e., the order of an adult's request) interact in a systematic and predictable way in determining children's extensions of animal labels.

In Study 1, we demonstrated that preschoolers use information about internal properties of animals to determine label assignment. Study 2 was designed to investigate, on the one hand, whether children use similar kinds of information in a domain in which such information may be less important or even irrelevant and, on the other hand, whether other kinds of properties play an "essential" role in determining category membership in a different domain. Study 2, then, looks at children's extensions of labels for artifacts. To the extent that children appropriately limit the role of internal properties to the animal domain, we can conclude that the results of Study 1 were not the product of a response bias to a particular set of instructions but rather reflected a deeper conceptual understanding. Furthermore, a finding that no type of property can lead children to believe two artifacts are of the same kind would support the notion that children do not hold essentialist beliefs about artifacts.

Study 2

It has been suggested that children conceptually distinguish between artifact categories and animal categories by treating only the latter as natural kinds (Gelman, 1988; Kalish, 1995; Keil, 1986). One proposed difference between animal and artifact categories is that whereas the former have an underlying nature that is believed to be essential to their identity, the latter do not (Schwartz, 1978). In fact, Atran (1995) argued that across cultures, children appreciate the distinction between the domains of living kinds and artifacts, based in part on the belief that living kinds have essences, whereas artifacts do not. An alternative possibility is that artifacts have essential properties that differ from those for animals.

To examine these alternative hypotheses, in Study 2, we tested whether telling children that two artifacts shared different types of properties would influence how they interpreted the relation between labels for artifacts in the same way as information about internal properties of animals affected children's interpretations of labels for animals. In Study 2, in addition to an internal and a superficial condition, we included a functional condition. As various researchers have argued (Keil, 1986; Nelson, 1995; Wierzbicka, 1984; but see Malt & Johnson, 1992), the function of an artifact may be criterial for determining its category membership. If children believe that functional properties are essential to determine the category membership of artifacts, then children in a functional condition should extend the familiar label to two artifacts and should reject a mutual exclusivity interpretation of the relation between two labels for different artifacts. In other words, children in a functional condition should respond differently from children in an internal or a superficial condition and similarly to the children in the internal-animal condition of Study 1. Alternatively, if function is not believed to be essential, there should be no condition effect, with children in all three conditions responding similarly to the children in the superficial-animal condition of Study 1. Specifically, they should stick to an assumption of mutual exclusivity. Study 2 also addressed whether children believe internal properties are essential in any domain or only for animals. Given that there was no effect of age in Study 1, only 4-year-olds were tested in Study 2.

Method

Participants

Forty-four 4-year-olds participated in this study (M = 4 years 7 months, range = 4 years 0 months to 4 years 11 months). There were
23 boys and 21 girls. An additional boy was dropped from the study for failing to complete the task. Thirteen children participated in the functional condition, 16 in the internal condition, and 15 in the superficial condition. The mean age of children in the three conditions did not differ significantly. All participants attended either one of two preschools affiliated with a midwestern university or a third preschool in the same city. They constituted a racially and ethnically mixed, middle-class sample. Twenty-five undergraduate students provided ratings of the stimuli.

**Design**

The experiment was a 3 (condition: functional, internal, superficial) x 2 (perceptual similarity: dissimilar, similar) x 2 (order: familiar first, novel first) design. Condition was a between-subjects variable; perceptual similarity and order were within-subjects variables.

**Stimuli**

The experimental stimuli for this experiment were 10 x 10-cm realistic color drawings of artifacts. Each child saw 10 sets of five pictures each. The pictures were arranged in an album in the same way as in Study 1. The order of pictures was randomly established and was the same for all children (guitar, stove, bus, thermometer, boots, car, piano, lawn mower, boat, axe). The album also included one warm-up set (a blue square, a red square, and a blue triangle) and two filler sets (butterflies, monkey, kangaroo, wolf, rabbit).

As in Study 1, the physical similarity between the two target pictures within each set was systematically varied. Each participant received five similar sets and five dissimilar sets (see Figures 3A and 3B). The perceptual similarity of the pictures was established by adult ratings. The mean score for dissimilars (on a scale of 1-7) was 2.76; the mean for similars was 5.63. This difference was significant, t(8) = 8.26, p < .0001 (see Table 4).

**Procedure**

In addition to the difference in items, the only other difference between Studies 1 and 2 involved the instructions. Children in the functional condition were told "This one [the atypical object] works in the same way as this one [the typical object]. You know, people use this one to do the same kinds of things as this other one." In the internal condition, the experimenter told children "This one [pointing to the atypical object] has the same insides as this one [pointing to the typical object]. You know, it has the same kind of parts and stuff inside as this one." In the superficial condition, the experimenter told children "This one [pointing to the atypical object] is the same size as this one [pointing to the typical object]. You know, people can buy this one in the same city as this other one."

The order of presentation and subsequent testing of the labels varied as in Study 1 to enable a direct comparison between the two studies and to replicate the pragmatic effect. For items 1, 3, 5, 7, and 9, children were first asked to identify referents of the familiar label and then of the novel label (Order 1). For items 2, 4, 6, 8, and 10, children were first asked about the novel label and then about the familiar label (Order 2).

**Results**

**Coding**

Children's responses were coded in the same way as in Study 1 (see Table 2 for an example).

**Label Extensions**

As in Study 1, a first set of analyses looked at how accurately children interpreted the familiar and novel labels. Arcsine trans-
Table 4  
Items Used in Study 2

<table>
<thead>
<tr>
<th>Category</th>
<th>Artificial labels used</th>
<th>M perceptual similarity between typical and atypical</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Similars</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lawn mower</td>
<td>Lore</td>
<td>5.2</td>
</tr>
<tr>
<td>Ax</td>
<td>Vit</td>
<td>6.5</td>
</tr>
<tr>
<td>Guitar</td>
<td>Zav</td>
<td>5.8</td>
</tr>
<tr>
<td>Bus</td>
<td>Kiv</td>
<td>4.6</td>
</tr>
<tr>
<td>Boots</td>
<td>Jops</td>
<td>6.0</td>
</tr>
<tr>
<td><strong>Dissimilars</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stove</td>
<td>Mef</td>
<td>2.5</td>
</tr>
<tr>
<td>Thermometer</td>
<td>Sap</td>
<td>2.4</td>
</tr>
<tr>
<td>Boat</td>
<td>Cak</td>
<td>3.0</td>
</tr>
<tr>
<td>Piano</td>
<td>Wug</td>
<td>3.2</td>
</tr>
<tr>
<td>Car</td>
<td>Zon</td>
<td>2.8</td>
</tr>
</tbody>
</table>

formed proportions of correct familiar and correct novel label extensions were entered into a repeated-measures ANOVA, with condition, perceptual similarity, and order as variables.

Condition did not have an effect on either measure. The only significant effect on correct interpretations of the familiar label was of order, $F(1, 39) = 35.88, p < .0001$. As in Study 1, children were more likely to extend the familiar label correctly to both artifacts when asked for the familiar label first ($M = 63\%$) than when asked for it second ($M = 35\%$). As for children's interpretation of the novel label, both order and perceptual similarity had significant effects. As in Study 1, children were more likely to learn the novel label when the artifacts were perceptually dissimilar ($M = 71\%$) than when they were similar ($M = 63\%$), $F(1, 39) = 12.96, p < .001$. Children were more likely to learn the novel label when asked for it second ($M = 69\%$) than when asked for it first ($M = 63\%$), $F(1, 39) = 7.67, p < .01$. None of the interactions on either measure were significant.

**Relation Between Labels**

For the main analyses on children's interpretations of the semantic relation between the labels, separate repeated-measures ANOVAs were conducted on the arcsine transformed proportion of answers of each response type (subordinate or mutual exclusivity) out of all responses children gave. In these analyses, gender and condition were included as between-subjects variables and perceptual similarity and order as within-subjects variables.

For subordinate responses, condition had no significant effect. As can be seen in Table 3, children in all three conditions made approximately the same number of subordinate responses, $F(2, 38) = 0.28, p > .70$. Curiously, girls made more subordinate interpretations ($M = 39\%$) than boys ($M = 16\%$), $F(1, 38) = 4.87, p < .05$. Perceptual similarity also had no significant effect, $F(1, 38) = 0.01, p > .90$. Finally, order had the expected effect on the proportion of subordinate interpretations. Children were more likely to make subordinate responses when they were questioned about the familiar label first (Order 1, $M = 39\%$) than when they were questioned about the novel label first (Order 2, $M = 15\%$), $F(1, 38) = 27.46, p < .0005$ (see Figure 4). None of the interactions were significant.

Condition also did not have a significant effect in terms of mutual exclusivity interpretations, $F(2, 38) = 0.20, p > .80$, and neither did perceptual similarity, $F(1, 38) = 0.85, p > .30$, though the effect of the latter was in the direction predicted (see Figure 4). Gender did not have a significant effect on this measure. Order, however, had a significant effect. Children were more likely to make mutual exclusivity interpretations when they were asked about the novel label first (Order 1, $M = 57\%$) than when they were asked about the familiar label first (Order 1, $M = 36\%$), $F(1, 38) = 27.67, p < .0005$. The interaction between order and perceptual similarity was also significant, $F(1, 38) = 4.37, p < .05$. Post hoc $t$ tests with adjusted $p$ values revealed a significant effect of order for similar pairs, $t(43) = 5.09, p < .05$, but not for dissimilar pairs, $t(43) = 2.26, p > .05$. None of the other interactions were significant.

ANOVA's on all the other response types revealed no significant effects.

![Figure 4](http://example.com/figure4.png)  
Study 2: Subordinate and mutual exclusivity responses by perceptual similarity and order.
Item Effects

Although condition did not have an effect on any of the measures, we decided to determine whether, for individual items, certain property information affected children's responses. In particular, we were interested in the possibility that for certain items for which internal parts are "constitutional" (e.g., parts that are crucial to the functioning or identity of the artifact), children in the internal condition would be more likely to make subordinate responses and less likely to make mutual exclusivity responses than children in either of the other two conditions.

For only 1 out of the 10 artifact pairs did children in the internal condition make substantially more subordinate and fewer mutual exclusivity responses than children in the superficial and functional conditions (substantially was defined as a difference of 10 percentage points or more on the percentage of responses of each type in the different conditions). Curiously, this item was guitar, an object for which the internal properties do not seem to be constitutional.

Comparisons Between Studies 1 and 2

To test more directly the domain generality of essentialist beliefs, we compared children's responses in Study 1 with children's responses in Study 2. Our main interest was to check whether internal property information in the domain of animals indeed had a special status, such that it was the only type of information to substantially lead children to believe the entities were of the same kind. For these analyses, we included only the 4-year-olds from Study 1 and only the children who participated in the internal or superficial condition in either study. This was done to have a balanced design for the comparison, controlling also for age. Thus we compared children's response patterns from four groups: internal-animal (n = 16), superficial-animal (n = 16), internal-artifact (n = 16), and superficial-artifact (n = 15).

We conducted two contrasts comparing children's responses in the internal-animal condition with the three other conditions (weights = 3, -1, -1, -1). For the first contrast, we used the number of subordinate interpretations as the dependent measure. We found that children in the internal-animal condition made more subordinate interpretations than children in the other three conditions, \( t(59) = 2.41, p < .05 \). In a second contrast, we used the number of mutual exclusivity interpretations as the dependent measure. Children in the internal-animal condition made fewer mutual exclusivity interpretations than children in the other three conditions, \( t(59) = -2.15, p < .05 \) (see Table 3).

These results were confirmed by analyses of children's individual response patterns. For these analyses, children were grouped according to whether they gave the same response type for more than half the items. Eight of the 16 children in the internal-animal condition gave subordinate responses for more than half the items, whereas 10 of 16 in the superficial-animal, 8 of 16 in the internal-artifact, and 7 of 15 in the superficial-artifact conditions did so. This pattern, however, did not reach significance, \( \chi^2(3, N = 63) = 4.70, p > .10 \).

In sum, children's response patterns in the internal condition of Study 1 differed from the pattern of responses of children from all three other groups. It was only when told that animals share internal properties that 4-year-olds overcame a mutual exclusivity bias, did not restrict the extension of a familiar label, and interpreted a novel label as a subordinate term. Children did not do so when told that two animals share superficial properties or that two artifacts share superficial, internal, or functional properties. This finding suggests that children do not hold essentialist beliefs in the domain of artifacts but do so in the domain of animals.

Discussion

Children in Study 2 were not affected by the kind of information they received about the properties that two artifacts shared. Across the three conditions, children interpreted artifact labels in strikingly similar ways. Children in the functional, internal, and superficial conditions were equally likely to extend the familiar label to both artifacts. Moreover, children in these three conditions made approximately the same number of mutual exclusivity interpretations and of subordinate interpretations. In sum, functional information did not differentially influence children's extensions of labels (see also Gathercole, Cramer, Somerville, & de Haan, 1995; Gentner, 1982; Smith et al., 1996), and it did not seem that children generalized the "essential" role of inside information from the domain of animals to the domain of artifacts. These findings are consistent with the claim that children do not have essentialist beliefs about artifacts.

However, before we can draw such a conclusion, at least one possibility has to be ruled out. Specifically, it may be that it is not so much that children do not have essentialist beliefs about artifacts but, rather, that these beliefs are not as specific for artifacts as they are for animals. In other words, it could be that children responded similarly to the property information in all three conditions in Study 2 because all three types of information indiscriminately probed children's essentialist beliefs about artifacts. Study 3 examines this possibility by providing a baseline condition in which no property information is given.

Somewhat surprising, the perceptual similarity between two artifacts did not affect children's interpretations of their labels. As mentioned before, Diesendruck and Shtz (1998) and Waxman and Senghas (1992) found that 2-year-olds were more likely to interpret labels for two artifacts as referring to mutually exclusive categories when the artifacts were perceptually dissimilar than when they were similar. Study 3 reexamines this
issue by more directly testing the interaction between perceptual similarity and domain.

A final point is that as in Study 1, children in Study 2 were influenced by the order in which the labels were tested. In Order 2, children first picked one of the objects as the referent of the novel label. When then asked for the familiar label, children's choice might have been somewhat limited to the other object. In other words, asking for the familiar label second might have highlighted to children the impending overlap in the extensions of the novel and the familiar labels—a violation of the mutual exclusivity bias—and thus led them to restrict the extension of the familiar label to only one object to maintain mutual exclusivity.

Study 3

Study 3 was designed to provide a baseline for children's responses. In this study, children were tested in a label-only condition in which no properties relating two entities were supplied. This study served several purposes. First, Study 3 provided an unbiased test of the interaction between perceptual similarity and domain. In Study 1, we found that the perceptual similarity between two animals significantly affected children's interpretation of the animals' labels. In Study 2, perceptual similarity did not have a significant effect on children's interpretation of labels for artifacts. By comparing the two domains in a label-only condition, we can examine the interaction between perceptual similarity and domain more directly.

Second, Study 3 allowed us to examine the direction of effects in Study 1. In Study 1, children in the internal condition made fewer mutually exclusive interpretations than children in the superficial condition. It could not be established, on the basis of these findings, whether (a) internal property information about animals drove children away from mutual exclusivity errors or (b) superficial property information drove them toward such responses. We interpreted the results as supporting the first conclusion, but a baseline condition is needed to rule out the latter account. By comparing children's responses in a label-only animal condition to their responses in the animal-property conditions, we could directly test these alternative hypotheses. If children in the label-only condition responded similarly to children in the superficial-animal condition, we could safely conclude that internal property information about animals indeed led children away from a mutual exclusivity interpretation of the labels.

Finally, in Study 2, the null effect of condition could be analogously interpreted in one of two ways. Either children did not hold essentialist beliefs about artifacts or they did, but such beliefs were equally probed by internal, superficial, and functional information. We could test these hypotheses by comparing children's responses in a label-only artifact condition with children's responses in the three conditions of Study 2. If children responded similarly in these four conditions, we could conclude that none of the types of property information probed children's essentialism.

Method

Participants

Thirty-one children participated in this study ($M = 4$ years 6 months, range = 4 years 0 months to 5 years 0 months). There were 11 boys and 20 girls who were about equally distributed between the two conditions. Three children were dropped from the study for failure to complete the task. Sixteen children participated in the label-only artifact condition, 15 in the label-only animal condition. There were no significant differences in the age of children from the different conditions. All participants attended preschools in a major city of the Southwest. Children were mostly from middle-class families from diverse races and ethnicities.

Design

The experiment was a 2 (domain: animals, artifacts) X 2 (perceptual similarity: dissimilar, similar) X 2 (order: familiar first, novel first) design. Domain was a between-subjects variable; perceptual similarity and order were within-subjects variables.

Stimuli

The materials used in this study were the same as those used in Study 1 for the domain of animals, and as those used in Study 2 for the domain of artifacts.

Procedure

The only difference between the procedure in this study and the procedures of Studies 1 and 2 is that in the present study, the experimenter did not provide children with any information about properties that the entities shared. The experimenter showed children the pair of target pictures in the book and immediately labeled them, in exactly the same form as was done in the previous two studies. The order of presentation and testing of the labels varied as in the previous studies.

Results

Children's responses were coded as in Studies 1 and 2. Given the results of the previous studies, the analyses focused exclusively on subordinate and mutual exclusivity responses. The first set of analyses tested for the effect of domain and its interaction with perceptual similarity on children's responses. The second set of analyses compared children's responses across the three studies. Preliminary analyses revealed no effect of gender on subordinate and mutual exclusivity responses, and it was therefore not included as a variable in the subsequent analyses.

Separate repeated-measures ANOVAs with domain as a between-subjects variable and perceptual similarity and order as within-subjects variables were conducted on the arcsine transformed proportions of subordinate and mutual exclusivity responses. Domain had a significant effect on the proportion of subordinate responses, $F(1, 29) = 5.87, p < .05$. As can be seen in Table 3, children made more subordinate responses for artifacts than for animals. The effect of perceptual similarity was not significant, $F(1, 29) = 0.01, p > .90$, and order again had a significant effect in the expected direction, $F(1, 29) = 15.68, p < .0005$. Most important, the interaction between perceptual similarity and domain was not significant, $F(1, 29) = 0.81, p > .30$.

With mutual exclusivity responses, domain again had a significant effect, $F(1, 29) = 16.44, p < .001$. Children were more likely to make mutual exclusivity responses for animal labels than for artifact labels (see Table 3). Perceptual similarity also had a significant effect, $F(1, 29) = 4.62, p < .05$, with children being more likely to make mutual exclusivity interpretations for
dissimilar pairs than for similar pairs. Finally, the effect of order replicated the findings of the previous studies, $F(1, 29) = 26.83, p < .0005$. Again, however, the interaction between perceptual similarity and domain was not significant, $F(1, 29) = 3.03, p > .05$. The other interactions were not significant.

To test for the effect of property information on children's responses, we conducted a series of planned contrasts. These contrasts compared 4-year-olds' responses in the experimental conditions in Study 1 with those in the label-only animal condition in Study 3, and their responses in the experimental conditions in Study 2 with those in the label-only artifact condition in Study 3.

Our interpretation of children's responses to animal labels was that internal information led children to overcome a mutual exclusivity bias. To test this hypothesis, two contrasts were set comparing children's responses in the internal condition with their responses in the superficial and label-only conditions, all for animals (i.e., weights = 2, -1, -1). We found that children in the internal condition made more subordinate responses, $t(44) = 3.15, p < .005$, and fewer mutual exclusivity responses, $t(44) = -3.27, p < .005$, than children in the two other conditions. In fact, in terms of both subordinate and mutual exclusivity responses, Scheffé multiple-comparisons tests revealed significant differences ($ps < .05$) between the internal and label-only conditions and no significant differences between the superficial and label-only conditions.

As for children's interpretation of artifact labels, we were interested in testing whether property information per se affected children's responses. For this purpose, two contrasts were conducted comparing the number of subordinate responses and the number of mutual exclusivity responses in the label-only condition with the corresponding numbers in the internal, superficial, and functional conditions (i.e., weights = -3, 1, 1, 1). These contrasts were not significant: for subordinate, $t(56) = -0.53, p > .5$; for mutual exclusivity, $t(56) = 1.97, p > .05$. In fact, Scheffé multiple-comparisons tests revealed no significant differences between any two groups on either of the two measures.

**Discussion**

The purpose of Study 3 was to clarify the findings of the previous two studies. The results of this study supported our earlier conclusions. First of all, the comparison of children's responses in the label-only conditions in the domains of animals and artifacts suggests that a priori, children were not more inclined to provide subordinate interpretations for the animal sets than for the artifact sets. In fact, the opposite was the case. Thus it seems that the finding that children in the internal-animal condition were more likely to make subordinate responses than children in any of the artifacts conditions (and in the superficial-animal condition) was not explained by children having some head start in the domain of animals. Rather, internal property information about animals indeed helped children make subordinate interpretations of novel labels and thus steered them away from mutual exclusivity interpretations.

This conclusion was further reinforced by the comparison of the label-only animal condition with the internal and superficial animal conditions of Study 1. Children made approximately the same number of subordinate and mutual exclusivity responses when told just labels for animals as when additionally told shared superficial properties for animals. However, children made significantly more subordinate and significantly fewer mutual exclusivity responses when told the animals shared internal properties than when simply told the animals' labels.

Study 3 also provided further evidence for the lack of essentialist beliefs in the domain of artifacts. It was not the case that children simply had a vaguer essentialist belief that encompassed internal properties, external properties, and functions. Children's responses when told that artifacts shared any of these kinds of properties were statistically indistinguishable from when they were simply provided with labels for artifacts. In other words, information about properties that two artifacts shared did not lead children to alter their interpretations of the artifacts' labels in significant ways.

Finally, the interaction between perceptual similarity and domain in Study 3 was not significant. This suggests that the finding of a significant effect of perceptual similarity in the domain of animals (Study 1), but not in the domain of artifacts (Study 2), should be interpreted with caution, as will be discussed in the General Discussion.

**Study 4**

The finding in Study 3 that children were more likely to make subordinate responses in the domain of artifacts than in the domain of animals could have resulted from different biases in the two domains or from differences in the initial familiarity of children with the items in the two domains. The purpose of Study 4 was to test how familiar children were with the items presented in the first three studies.

In this study, we simply presented children with pictures of the typical and atypical exemplars of each category and asked them to point to referents of the familiar labels used in the previous studies. We were interested in seeing how likely children were to extend the familiar label to both exemplars.

**Method**

**Participants**

Twelve 4-year-olds participated in this study ($M = 4$ years 6 months, range = 3 years 9 months to 4 years 10 months). There were 5 boys and 7 girls. All children attended a preschool in a major city of the Southwest. Children were mostly from middle-class families and from diverse ethnic backgrounds.

**Design**

The experiment was a 2 (domain: animals, artifacts) $\times$ 2 (perceptual similarity: dissimilar, similar) $\times$ 2 (order of domain: animals first, artifacts first) design. Domain and perceptual similarity were within-subjects variables; order of domain was a between-subjects variable.

---

3 We thank two anonymous reviewers for suggesting this follow-up study.
Stimuli

The materials used in this study were the same as those used in the previous studies. In contrast to the previous studies, however, all 20 items (10 animals and 10 artifacts) were arranged in a single album, and only the "test" sets (i.e., the pictures in the right-hand pages in the previous studies) were presented. Thus, children saw 20 triads of pictures, each triad consisting of pictures of two exemplars of an animal or an artifact category plus a distractor animal or artifact. As in the previous studies, 5 of the pairs of exemplars in each domain were perceptually similar, and the other 5 pairs were dissimilar. Thus, altogether there were 10 similar pairs and 10 dissimilar pairs. The order in which the sets of animal pictures were arranged in the album was the same as in Studies 2 and 3. The order in which the sets of artifact pictures were arranged in the album was the same as in Studies 2 and 3.

Procedure

The experimenter presented children one page at a time and asked them to identify referents of the familiar labels used in the previous studies (e.g., "Can you show me a squirrel?" "Is there another squirrel?" and so on, until the child said "no"). For methodological reasons, we counterbalanced, between subjects, the order in which the sets of pictures of each domain were presented. Thus, for half of the children, the experimenter first showed the 10 sets of animal triads and then the 10 sets of artifact triads; for the other half, the experimenter first showed the artifact triads and then the animal triads.

Results and Discussion

Our main interest was whether children would pick both the typical and atypical exemplars of each category as referents of the familiar label. Thus, children's responses were simply coded as a 1 if they picked both exemplars or a 0 if they did not. A repeated-measures ANOVA was conducted on the number of "full" extensions that children made. Domain and perceptual similarity were entered as within-subjects variables and gender and order as between-subjects variables.

Children were more likely to extend the familiar label to the two exemplars for artifacts (M = 8.75) than for animals (M = 5.33), F(1, 8) = 61.94, p < .0001. Children were also more likely to extend the familiar label to both exemplars when the atypical exemplar was perceptually similar to the typical exemplar (M = 9.75) than when it was dissimilar (M = 4.33), F(1, 8) = 94.28, p < .0005. The interaction between domain and perceptual similarity was also significant, F(1, 8) = 121.88, p < .0001. Post hoc tests revealed a significant effect of domain for dissimilar items (M = 3.75 for artifacts, M = 0.58 for animals, both out of 5), t(11) = 10.65, p < .001, but not for similar items (M = 5.00 for artifacts, M = 4.75 for animals, both out of 5), t(11) = 1.91, p > .05. There were no effects involving gender or order of domains.

The finding of an effect of domain complicates our comparison of children's responses across the two domains. Specifically, for dissimilar items, it might be that to begin with, children in the two domains were confronted with different tasks. To control for this possible confound in our comparison of the domains, we reanalyzed the contrasts comparing children's responses in Studies 1 and 2 by including only their responses for similar items, which yielded equivalent performance across domains in Study 4. For these items, children's initial predispositions in the two domains were the same.

We conducted the same two contrasts done earlier: one comparing the mean number of subordinate responses and the other, the mean number of mutual exclusivity responses. In both contrasts, children's responses in the internal-animal condition from Study 1 were compared with the children's responses in the superficial-animal condition from Study 1, internal-artifact condition, and superficial-artifact condition, these last two from Study 2 (i.e., weights = 3, -1, -1, -1).

The results were equivalent to those reported earlier. Children were more likely to make subordinate responses in the internal-animal condition (M = 2.87, out of 5) than in any of the other three conditions, r(59) = 2.53, p < .02 (means for superficial-animal, internal-artifact, superficial-artifact were 1.81, 1.37, 1.67, respectively). Conversely, children were less likely to make mutual exclusivity responses in the internal-animal condition (M = 1.06) than in any of the other three conditions, r(59) = -2.52, p < .02 (means for superficial-animal, internal-artifact, superficial-artifact were 2.56, 2.25, 2.13, respectively).

In sum, even when controlling for children's initial extensions of the familiar labels, internal property information about animals clearly led children to accept a common label for two animals and to reject a mutual exclusivity interpretation of labels for animals. It is also interesting to notice that because children's initial extensions of the familiar animal labels were narrower than those of artifact labels (on dissimilar sets), it is all the more remarkable that children's subsequent extensions of the familiar label were wider in the internal-animal condition than in any of the artifacts conditions (as manifested in terms of both more subordinate responses and fewer mutual exclusivity responses). In other words, domain differences obtained in the experimental conditions of Studies 1 and 2 are the reverse of the domain differences displayed in children's initial interpretations of the familiar labels (Study 4). We suggest that this resulted from the fact that while the mutual exclusivity bias led children to restrict the extension of the familiar label in the domain of artifacts and when only labels or superficial properties were provided for animals, internal property information about animals—presumably by tapping into children's essentialist beliefs—led them to overcome mutual exclusivity and not to restrict the extension of the familiar label.

General Discussion

Children who were told that two animals shared internal properties were more likely to conceive of the animals as being of the same kind than children who were told that the animals shared superficial properties. This was manifested in mainly two ways: children in the internal condition were more likely to extend the familiar label to both animals and were less likely to treat labels for these animals as referring to mutually exclusive categories than children in the superficial condition. The superficial condition served as an important control in that as with the internal condition, it described properties that the animals shared. Thus, the finding that children in the internal condition made fewer mutually exclusive interpretations cannot be based on a general strategy to treat animals with common properties.
as having the same name. Rather, children treated animals that share internal, in contrast to superficial, properties as more likely to be taxonomically related and thus to share a common label. The label-only condition (Study 3) also served as an important control by showing that it was indeed internal information about animals that significantly altered children’s unbiased responses.

Various researchers have argued that children engage in essentialist reasoning in categorizing animals, in making inferences about properties of animals, or in identifying animals (e.g., Gelman et al., 1994; Gelman & Wellman, 1991). What is original about the current findings is that they demonstrate how essentialist beliefs underlie children’s decisions about word meaning. As various students of language and cognitive development have pointed out, the ties between semantic and conceptual development are not simple and direct (Keil, 1989; Soja et al., 1992; Waxman, 1991). Remarkably, nonetheless, children in the present studies quite straightforwardly concluded that if two animals have common essential properties, then they should have a common name. As pointed out in the introduction, this inference is the most direct implication of essentialist reasoning.

Study 2 explored the role of essentialist beliefs on the extension of labels for artifacts. One could have expected the function of an object to be a likely candidate for determining its category membership (Keil, 1986; Nelson, 1995; Putnam, 1975). To use Medin and Ortony’s (1989) term, one could have argued that function serves as a ‘‘essence placeholder’’ in the domain of artifacts in the same way as internal properties are claimed to serve as essence placeholders in the domain of animals. Our findings do not support this contention. Children were not affected by the sort of property that two artifacts shared, neither in terms of how they extended the familiar label nor in terms of how they interpreted the relation between labels. Methodologically, these findings also serve as another control for our results with animals. As our cross-study comparisons indicate, internal properties influenced children’s interpretation patterns only in the domain of animals. Children’s responses in all other conditions were statistically indistinguishable.

These findings with artifacts support claims about the domain specificity of essentialist beliefs (e.g., Atran, 1995), though two alternative accounts are tenable. First, it is possible that the functional information provided in Study 2 was too abstract and not sufficiently emphasized to trigger children’s essentialist reasoning. In Nelson (1995), for instance, preschoolers practiced and had much exposure to the specific function of a single novel artifact. Subsequently, when asked to extend the name of that novel artifact, children did so on the basis of whether objects could perform the novel artifact’s function. Second, it may be the case that essentialist beliefs about artifacts are geared toward yet a different set of properties—for instance, the original intention of an artifact’s creator (Bloom, 1996).

Nonetheless, converging evidence from different perspectives ratifies the conclusion that children as well as adults do not have essentialist beliefs about artifacts. In contrast to animal categories, children seem to treat certain artificial categories (e.g., tools) as social conventions—that is, groupings not objectively defined by nature (Kalish, 1995). Work with adults shows that though function is important for classifying artifacts, it does not supply a conceptual core (Malt & Johnson, 1992). Furthermore, students of linguistics and philosophy have argued that terms for “purely functional” artifacts are different from terms for animals in that they do not designate kinds of things (Wierzbicka, 1984) and do not capture common hidden structures (Schwartz, 1978). Consistent with these claims, it has been argued that across cultures, children distinguish between living kinds and artifacts in that only the former are presumed to have underlying causal essences (Atran, 1995).

It is interesting that we also found that the two domains differed somewhat in terms of the importance of perceptual similarity. The degree to which two entities looked alike affected children’s interpretations of labels when the entities labeled were animals (Study 1) but not when they were artifacts (Study 2). As expected, labels for similar looking animals were less likely to be interpreted as denoting mutually exclusive categories than labels referring to dissimilar animals. It is important, though, that the interaction between domain and perceptual similarity was not significant (Study 3). Given our findings in Study 4, however, we have to interpret these findings with caution. Specifically, Study 4 revealed an interaction between perceptual similarity and domain in terms of children’s extensions of the familiar labels. The difference between similar and dissimilar items in the extent to which children extended the familiar label to both entities was larger for animals than for artifacts. It is possible that this difference in children’s initial predispositions accounts for the differential effect of perceptual similarity in the two domains that was found in Studies 1 and 2. Further studies are needed to test this interaction more directly.

The finding that perceptual similarity had no effect on 4-year-olds’ interpretations of labels for artifacts also seems to contradict findings about the significance of perceptual similarity in toddlers’ extensions of object labels (Diesendruck & Shatz, 1998; Waxman & Senghas, 1992). One possible reconciliation of these studies is that there are developmental changes in the relevance of perceptual similarity, specifically in terms of children’s growing domain-specific knowledge (for a discussion, see Keil, 1995; Jones & Smith, 1993; Medin, Goldstone, & Gentner, 1993). The present studies were not designed to investigate this question, though it is an important one to pursue.

As expected, across studies, we found that the order of questioning affected children’s responses. We had hypothesized that an interaction between a mutual exclusivity assumption and pragmatic considerations would lead to such effects. As revealed by our analyses of how accurately children extended the familiar and novel labels, order primarily affected how children extended the familiar label and not so much how they extended the novel one. In Order 1, children were asked to pick referents for the familiar label first, and thus their choices were unconstrained. That is, they had not yet committed to one of the exemplars as a referent of another label; thus, they were free to follow their initial intuitions and to pick both exemplars as referents of the familiar label. By subsequently picking one of the exemplars again as the referent of the novel label, children violated mutual exclusivity and, in our coding scheme, instantiated a subordinate relation between the labels. In Order 2, however, children were first asked to pick referents for the novel label; thus, by the time they were asked for referents of the familiar label, their choices were compromised. Namely, by picking again the referent of
the novel label as a referent of the familiar label, they would be clearly violating mutual exclusivity. To avoid doing so, children, in these cases, corrected their extension of the familiar label and picked only the unlabeled exemplar as the referent of the familiar label (see Merriman & Stevenson, 1997, for somewhat comparable findings).

Finally, whereas the studies mostly focused on how conceptual, perceptual, and pragmatic factors influence children's label extension, they also provide strong evidence for Markman's (1989) mutual exclusivity bias. Mutual exclusivity responses were the predominant response type overall (see also Gelman et al., 1989, for similar findings), accounting for approximately 45% of children's responses. This is especially remarkable, given that the labeling instructions implied a hierarchical relation between the two labels and that in most conditions children were told about properties two entities shared rather than properties that differentiated the entities. The findings further confirm, though, that children did not treat the mutual exclusivity bias as an all-or-none hard constraint but rather as a default assumption that could be overridden by other factors (Woodward & Markman, 1991). The most direct evidence in support of this conclusion is that children made significantly fewer mutual exclusivity responses when told that two animals share internal properties than when told that (a) animals share superficial properties, (b) artifacts share internal, superficial, or functional properties, or (c) no additional information.

Children's initial extensions of the familiar labels (Study 4) are also revealing about the status of mutual exclusivity as a default assumption. For artifacts, children initially extended the familiar label to both objects on 87% of the items. After being taught a novel label for one of the objects (Study 2), however, children extended the familiar label to both objects on only 50% of the items. For animals, 4-year-olds initially extended the familiar label to both animal exemplars on just over 50% of the items (mostly the similar ones). After being taught a novel subordinate label for one of the animals and being told that the animals shared superficial properties (Study 1), 4-year-olds' full extensions of the familiar label dropped to 40% of the items. In other words, simply being taught a subordinate label for one of the exemplars led children to substantially restrict their extension of a familiar label to preserve mutual exclusivity (see also Merriman & Bowman, 1989; Merriman & Stevenson, 1997). In contrast, when in addition to being taught a subordinate label for one of the animals, children were told that the animals share internal properties, 4-year-olds' full extensions of the familiar labels rose to 70% of the items. In sum, by default, children assumed that the labels picked out mutually exclusive categories, but the essentialist information—as well as perceptual similarity and being asked for the familiar label first—led them to override this assumption.

The pattern of results described above is also informative about the role of property information in convincing children that two entities were of the same kind. One could have argued that children came into the label-learning situation already knowing that the similar-looking animals were of the same kind, and thus for these items, the internal property information was redundant. In contrast, our findings indicate that the internal property information was crucial in determining children's interpretations of the labels in the experimental conditions. Apparently, although children initially assumed that the similar-looking animals were both named with the familiar label, the new label coupled with a mutual exclusivity assumption put them into a conflict. Participants could have resolved this conflict by restricting the scope of the familiar word or by appropriately interpreting the new word as labeling a subtype. The findings suggest that children in the animal condition required the internal property information to maintain their conviction that similar-looking animals were indeed of the same kind. Without this information, children relinquished their initial intuitions about the extension of the familiar labels and fell back into a mutual exclusivity assumption—just as they did in all but the internal property condition of Study 1.

In conclusion, the present findings provide evidence supporting the view that children's naming does not occur in a conceptual and pragmatic vacuum. Any account of how conceptual information underlies children's naming is, by the nature of the problem, an interactionist account that involves consideration of children's naming biases. The same is true of the lexical constraints account, which involves consideration of the specific conceptual, perceptual, and pragmatic context in which words are taught. The present studies are a step toward a comprehensive account of word learning. In general, the studies expose specific ties between conceptual knowledge, pragmatic considerations, attention to perceptual information, and lexical assumptions in children's extensions of object labels. More specifically, the studies reveal the centrality and domain specificity of essentialist beliefs in children's concepts and how these beliefs underlie children's naming.

References


CONCEPTUAL AND LINGUISTIC BIASES


Received February 4, 1997
Revision received December 1, 1997
Accepted December 1, 1997