Two-year-olds’ recognition of hierarchies
Evidence from their interpretation of the semantic relation between object labels

Gil Diesendrucka,*, Marilyn Shatzb

aDepartment of Psychology, Bar-Ilan University, Ramat Gan 52900, Israel
bUniversity of Michigan, Ann Arbor, MI, USA

Received 1 March 2000; accepted 1 May 2001

Abstract

The present studies investigated 2-year-olds’ recognition of hierarchies by examining under various conditions the relative frequency with which children interpreted novel and familiar labels for objects as mutually exclusive. Two-year-olds were taught novel labels in one of three ways: (a) inclusive input: “[novel label] is a kind of [familiar label]”; (b) exclusive input: “[novel label] is not a [familiar label]”; or (c) (Study 2) no relational input. The referents of the novel and familiar labels were taxonomically either: (a) strongly related (e.g., a fighter airplane and a passenger airplane), or (b) weakly related (e.g., a paint brush and a toothbrush). Children were less likely to interpret the labels as picking out mutually exclusive categories when: (a) the labels were introduced with the inclusive input; and (b) the referents of the labels were taxonomically strongly related. This modulation of mutual exclusivity interpretations in response to the various hierarchical relations instantiated in the stimuli and input provides evidence for 2-year-olds’ emerging capacity to recognize hierarchical relations. © 2001 Elsevier Science Inc. All rights reserved.

1. Introduction

One of the fundamental characteristics of adult cognition is the capacity to organize knowledge hierarchically, with categories at different levels of
inclusiveness. Contrary to Piaget’s beliefs that young children lacked this capacity (Inhelder & Piaget, 1964), recent research on word learning reveals that under certain facilitating circumstances (e.g., when subordinate-level distinctions are highlighted), even 3-year-olds can categorize objects at different hierarchical levels (Callanan, 1989; Gelman, Wilcox, & Clark, 1989; Mervis, Johnson, & Mervis, 1994; Waxman & Hatch, 1992; Waxman, Shipley, & Shepperson, 1991). Whether even younger children have such capacities, however, is less clear.

There are conflicting findings on whether or not 2-year-olds are willing to apply a superordinate or subordinate label and a basic-level label to the same object (e.g., “toy” for a doll, or “animal” for a dog) (Anglin, 1977; Callanan & Markman, 1982; cf. Macnamara, 1982). Taylor and Gelman (1989) found that 2-year-olds were willing to extend a novel label (“fep”) to members of a subcategory (green cars) and to continue extending the familiar label (“car”) to all exemplars of the basic-level category (green and white cars). This pattern of extensions is consistent with children interpreting the novel label as subordinate to the familiar label. However, as Taylor and Gelman concluded, their task did not require children to “keep in mind both levels of a hierarchy at once” (p. 634). Similarly, Blewitt (1994) found that although 2-year-olds readily applied a basic-level label (e.g., “monkey”) and a superordinate label (e.g., “animal”) to the same object when asked separately for their referents, they seemed confused when asked about the two labels simultaneously (e.g., “Can it really be a monkey and an animal at the same time?”). Finally, Merriman (1986) found that 2-year-olds had difficulty understanding the relation between two labels at different hierarchical levels as one of inclusion (see also Merriman & Stevenson, 1997).

The purpose of the present studies was to shed light on these conflicting findings regarding 2-year-olds’ hierarchical understanding by investigating more precisely the circumstances under which 2-year-olds recognize that objects have an inclusion relation. More specifically, we investigate whether 2-year-olds are sensitive to the degree of taxonomic relatedness between objects, and to the type of hierarchical relation between objects implied by different linguistic input.

1.1. Taxonomic relatedness

An integral part of a mature understanding of hierarchies is the recognition that objects of the same kind are likely to be referred to by the same name (Blewitt, 1994). One of the main goals of the present study was to investigate whether 2-year-olds have this understanding. Taylor and Gelman’s (1989) study mentioned above may provide a clue on this issue. They found that 2-year-olds were quite likely to construct an inclusion relation between a novel and a familiar label. An aspect of their procedure that possibly helped children construe an inclusion relation between the labels was that the new objects were actual
exemplars of subordinate categories included in a familiar basic-level category (e.g., a beach ball and a soccer ball). In other words, perhaps 2-year-olds recognize that when objects are of the same basic-level kind, they are likely to share a common name. If that is the case, then we would predict that when objects are less strongly related (e.g., a passenger airplane and a glider), 2-year-olds will not expect them to share a common name. The present studies address this possibility by systematically varying the degree to which objects in a pair were of the same kind and analyzing its effect on children’s interpretation of labels for the objects.

In the current studies, children are presented pairs of objects that were considered, or not, by adults to be “very much the same kind of thing at the basic level” \(^1\). If 2-year-olds are sensitive to the degree of taxonomic relatedness between objects, they should be less likely to interpret labels for objects as mutually exclusive when the referents of the labels are very much the same kind of thing than when the referents are not at all the same kind of thing. This response pattern would indicate that 2-year-olds are sensitive to the hierarchical relations among object categories. Study 2 tests this capacity more directly by explicitly asking children if an object can have two names.

1.2. Linguistic input

Various researchers have noticed that when teaching new superordinate or subordinate labels to their children, mothers commonly use the phrasing “kind of” to point out the relation between the new label and a familiar basic-level label (Callanan, 1985; Hall, 1994). Studies with preschoolers found that this type of input indeed helped children acquire an inclusion relation between a new and a familiar label (Callanan, 1989; Gottfried & Tonks, 1996). Experimental work with toddlers, however, has provided less conclusive evidence on the facilitatory effect of inclusive input. In particular, Diesendruck and Shatz (1997) found that 2-year-olds who were told that a [new label] was a “kind of” [familiar label] were no less likely to interpret the labels as mutually exclusive than were children who were told that a [new label] was “not a” [familiar label].

The present studies also test whether 2-year-olds can understand the semantic relations expressed in different types of input. In Study 1, children are exposed to input conditions identical to those used by Diesendruck and Shatz (1997), but the overall procedure was modified to make it simpler for 2-year-olds. In Study 2, children are tested in a baseline condition in which no relational input is provided. This is done so as to assess children’s unprimed interpretation of the

\(^1\) Note that the goal of the paper was not to investigate how children determine whether or not objects are very much the same kind, but rather whether children are sensitive to the categorization of objects into kinds when deciding how to interpret the relation between names for objects. For that purpose, we simply operationalized the variable of degree of taxonomic relatedness by asking adults to judge object pairs on this criterion.
relation between the labels. If 2-year-olds are capable of understanding the hierarchical implication of the Inclusive and the Exclusive instructions, then we should find more mutual exclusivity interpretations of the labels in the latter than in the former condition.

1.3. Methodological issues

The most straightforward way to assess whether children understand hierarchical relations in naming tasks such as the present one is to analyze how often they interpret labels as denoting subordinate categories (Gelman et al., 1989; Mervis et al., 1994; Waxman et al., 1991). This strategy, however, poses a few concerns, especially for studying very young children. A general concern is that for children to be credited with an understanding of hierarchical relations, they have to generate explicitly an inclusion relation. In other words, the criterion for crediting children with an understanding of hierarchies is very stringent. A second concern is that it is very hard to establish whether children’s subordinate responses truly reflect an inclusion relation or some other kind of overlapping relation that does not involve hierarchical understanding (e.g., categories with common members such as “dog” and “pet”).

With these concerns in mind, and given our interest in children’s recognition (as opposed to generation) of hierarchies, we designed the present studies in a manner that would allow children to manifest an understanding of hierarchies less directly. Rather than relying exclusively on children to manifest an understanding of hierarchical relations directly through subordinate responses, we incorporate into the stimuli distinctions relevant to that capacity. We then evaluate children’s sensitivity to these distinctions with an alternative measure of a response type that should be more accessible to them than subordinate responses. In particular, we imply different hierarchical relations between objects by varying the degree of taxonomic relatedness between objects and the linguistic input by which object labels are introduced. The question then becomes whether 2-year-olds can discriminate among these different implications. To evaluate this question, we examine the frequency with which children respond to such stimulus differences with mutual exclusivity responses. On the one hand, to the extent that children do not discriminate among our conditions with varying amounts of mutual exclusivity responses, they would reveal a lack of understanding of hierarchical relations. On the other hand, variation in the frequency of mutual exclusivity by condition would reveal sensitivity to the hierarchical information.

Our use of mutual exclusivity response frequencies as an indicator of sensitivity to hierarchical information is based on two arguments. First, as noted above, subordinate responses, by themselves cannot be taken as decisive evidence for an understanding of inclusion relations; variation in mutual exclusivity responses as a function of hierarchical information is hard to interpret as anything other than sensitivity to such information and should at least be used along with the measure of subordinate responses. Second, there is
considerable evidence that mutual exclusivity responses are readily accessible to children (Markman, 1989; Merriman & Bowman, 1989). For example, 2-year-olds restrict or correct the extension of a familiar label upon learning a novel label, so that the extensions of the two labels do not overlap (Liittschwager & Markman, 1994; Merriman & Stevenson, 1997). Moreover, children have also demonstrated the ability to override the mutual exclusivity response tendency (Woodward & Markman, 1991). For instance, it has been shown that 2-year-olds’ tendency to interpret object labels as mutually exclusive is weaker when the objects are highly perceptually similar (Diesendruck & Shatz, 1997; Waxman & Senghas, 1992).

In addition to relying primarily on the variation in mutual exclusivity responses by condition, we also examine the frequency of subordinate responses. Despite the problems connected to them, it is of interest to consider children’s ability to generate them in the different conditions. The modulation of mutual exclusivity interpretations and a potentially complementary modulation of subordinate interpretations, in response to the various hierarchical relations instantiated in the stimuli and input, serve together as a way to assess more accurately 2-year-olds’ recognition of hierarchical relations. We hypothesize that mutual exclusivity interpretations will be infrequent, and subordinate interpretations common, when objects are strongly related taxonomically or when the input is inclusive.

2. Study 1

2.1. Method

2.1.1. Subjects

Forty 2-year-olds participated in this study ($M = 2$ years 4 months, range = 2 years 1 month to 2 years 9 months). There were 14 girls and 26 boys, from varied socio-economic and ethnic backgrounds. Thirty-five of the children were tested in a language development laboratory. The other five children were tested in a quiet area in a half-day preschool affiliated with a major midwestern university (two were assigned to the Inclusive input/Strongly related pair condition, and the remaining three children were each assigned to one of the other three conditions). Parental consent was obtained prior to testing all children. In addition, 20 undergraduate students provided ratings of the stimuli.

2.1.2. Design

The study had a 2 (degree of taxonomic relatedness: Strongly related or Weakly related) × 2 (input: Inclusive or Exclusive) experimental design. Both degree of taxonomic relatedness and input were between-subjects factors, resulting in four experimental conditions: (a) Inclusive input/Strongly related
pair; (b) Inclusive input/Weakly related pair; (c) Exclusive input/Strongly related pair; and (d) Exclusive input/Weakly related pair. There were 10 children in each condition, with three girls in each of the Strongly related conditions, and four girls in each of the Weakly related conditions. The mean age of children in the four conditions did not differ statistically.

2.1.3. Stimuli

To assess the degree of taxonomic relatedness among pairs of objects, 10 adult subjects were instructed to judge on a seven-point scale the degree to which two objects were “the same kind of thing” at a basic level (1 = not at all, 7 = very much). Specifically, they were told to consider whether two objects belonged to the same category in the same way as a sedan and a race car can be said to belong to the same category, but a train and a car do not. Altogether, adults rated 13 pairs of objects.

To assure that any difference we found in children’s interpretations of labels for Strongly related and Weakly related objects would be due primarily to the difference in the degree of taxonomic relatedness between the objects, we controlled for the degree of perceptual similarity between objects in each of the 13 pairs. An additional group of 10 adults was instructed to judge on a seven-point scale the degree to which the two objects in a pair looked alike, taking into account shape, size, complexity, and individual features (1 = not at all similar, 7 = extremely similar).

The criteria for selecting eight pairs for use with the children were that four Strongly related pairs should be significantly different from Weakly related pairs in terms of degree of taxonomic relatedness, but not different in terms of perceptual similarity. For degree of taxonomic relatedness, the mean score for Strongly related pairs was 5.6, and the mean score for Weakly related pairs was 2.7 [t(6) = 4.76, P < .005]. In terms of perceptual similarity, the mean score for Strongly related pairs was 4.9, and the mean for Weakly related pairs was 4.3 [t(6) = 1.68, P > .14]. The set of object pairs used in this study, together with the adult ratings on degree of taxonomic relatedness and perceptual similarity, is presented in Table 1. Table 1 also lists the distractor items presented with each target pair. The distractors were not from the same superordinate category as the target objects.

2.1.4. Procedure

Children were tested individually by a female experimenter in one 15–25 min session. The experimenter sat with the child at a table and said they were going to play with some toys she had in a box. The first two sets of objects presented to children were used for practice trials. The goal of these trials was to familiarize children with the procedure, in particular, to assure that the toddlers were comfortable with answering “no” to any particular question and that they were willing to pick more than one object in response to a single label. In one of the practice trials, the experimenter took three objects out of the
box (a bicycle, a helmet, and a pencil) without labeling them, and after a brief period of play with the objects, the experimenter asked the child to identify them (e.g., “Can you show me a bicycle?”, and then “Is there another bicycle?”). Children were corrected when they answered wrongly, and the experimenter made it clear that the child could say “no” in response to her questions. The experimenter next showed three other objects (two watches and a hanger), and again asked the child to identify the objects. In this second trial, children were encouraged to pick the two watches in response to the experimenter’s requests for a “watch.”

Following the practice trials, the experimenter conducted the four test trials, one for each of the four Strongly related or four Weakly related pairs of objects, depending on the child’s condition. The four sets of objects were presented in a random order across subjects. On each trial, the experimenter showed the child two distractor objects and a pair of target objects. The experimenter picked one object at a time from a box and labeled it while presenting it to the child. The first and last objects presented in each trial were always distractors, while the second and third objects were always target objects. The experimenter labeled one of the target objects using its familiar basic-level label (see Table 1). We refer to this as Object A. The other target object was given a novel label. We refer to this as Object B. A novel nonsense label was used to introduce Object B so as to assure that all children were unfamiliar with the labels being taught. These labels were the same in the four conditions (fep, dax, wug, and jop), and were assigned to the same target pair across subjects (see Table 1). Objects A and B were the same for all the children in each of the degree of taxonomic relatedness conditions, but the order in which they were introduced was counterbalanced between and within subjects.

Table 1
Stimulus pairs for Studies 1 and 2: objects, labels, and mean adult ratings of degree of taxonomic relatedness and perceptual similarity

<table>
<thead>
<tr>
<th>Object A (familiar label)</th>
<th>Object B (novel label)</th>
<th>TR</th>
<th>PS</th>
<th>Distractors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strongly related</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shoe (shoe)</td>
<td>sandal (fep)</td>
<td>4.9</td>
<td>4.2</td>
<td>truck, plate</td>
</tr>
<tr>
<td>Passenger plane (airplane)</td>
<td>fighter plane (dax)</td>
<td>5.4</td>
<td>5.4</td>
<td>wrench, funnel</td>
</tr>
<tr>
<td>Sailboat (boat)</td>
<td>motorboat (wug)</td>
<td>5.5</td>
<td>4.8</td>
<td>goggles, triangle</td>
</tr>
<tr>
<td>Hat (hat)</td>
<td>cap (jop)</td>
<td>6.4</td>
<td>5.3</td>
<td>football, motorcycle</td>
</tr>
<tr>
<td><strong>Weakly related</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Milk bottle (bottle)</td>
<td>soda bottle (fep)</td>
<td>3.7</td>
<td>5.0</td>
<td>van, hammer</td>
</tr>
<tr>
<td>Paint brush (brush)</td>
<td>toothbrush (dax)</td>
<td>1.4</td>
<td>4.0</td>
<td>glasses, ball</td>
</tr>
<tr>
<td>Purse (bag)</td>
<td>paper bag (wug)</td>
<td>3.3</td>
<td>4.1</td>
<td>chair, spoon</td>
</tr>
<tr>
<td>Ring (ring)</td>
<td>plastic ring (jop)</td>
<td>2.2</td>
<td>4.2</td>
<td>flower, glass</td>
</tr>
</tbody>
</table>

TR = mean for degree of taxonomic relatedness between objects in a pair; PS = mean for perceptual similarity between objects in a pair. The following items were not toys: shoe, sandal, paint brush, toothbrush, ring, funnel, and spoon.
For instance, the experimenter showed a child in the Strongly related condition a wrench (a distractor), said “here’s a wrench,” and put it on the table. She then picked up a toy passenger airplane (Object A) and said “here’s an airplane.” The other target object — Object B — was then introduced and labeled according to the child’s input condition. Specifically, about a toy fighter airplane: “this is a dax. It’s a kind of [Inclusive input] OR not an [Exclusive input] airplane, it’s a dax. A dax is a kind of OR not an airplane. This is a dax.” Finally, a second distractor was introduced (“here’s a funnel”).

Child and experimenter played with the objects for 2–3 min and the experimenter then said that it was time to clean up the objects so as to get other objects from the box. First though, she asked the child to identify the target objects. For instance, “Can you show me an airplane?” and if the child pointed to an object, “Is there another airplane?”, and so on until the child said “no.” The experimenter then asked about the novel label: “Can you show me a dax?” and “Is there another dax?”. The order of the questions followed the order in which the objects were introduced. Following the requests for the target labels, the experimenter returned the toys to the box, and introduced the next set of objects.

2.2. Results and discussion

2.2.1. Scoring

Children’s responses were coded according to the semantic relation they established between the novel label and the familiar label. (a) Mutually exclusive: the child picked only Object A (e.g., passenger airplane) as the referent of the familiar label (e.g., “airplane”) and only Object B (e.g., fighter airplane) as the referent of the novel label (e.g., “dax”). This response is consistent with a mutual exclusivity bias. (b) Subordinate: the child picked both Objects A and B (passenger airplane and fighter airplane) as the referents of the familiar label (“airplane”) but only Object B (fighter airplane) as the referent of the novel label (“dax”). This response is consistent with an inclusion relation between the labels, though as pointed out in the Introduction, it could also reflect other overlapping relations. (c) Synonymy: the child picked both objects as referents of both labels. (d) Familiar-only: the child did not pick Object B as the referent of the novel label, and picked both objects as referents of the familiar label.\(^2\) Other responses that did not fall into any of the above categories included: (e1) the child picked

\(^2\) A Familiar-only response is ambiguous: It could reflect what Merriman and Bowman (1989) referred to as the “rejection effect,” with children unwilling to grant a new name to an object similar in kind to a previously named object. Alternatively, it could result from children not having attended to — and thus learned — the novel label. In either case, we would not want to count such responses together with the ones children give when they clearly select different objects for the different labels, i.e., clear mutual exclusivity responses. Therefore, we put Familiar-only responses in a separate category.
Object A as the referent of the familiar label and one of the distractors as the referent of the novel label; and (e2) the child picked Object B and one of the distractors as referents of the novel label.

Separate analyses of variance (ANOVAs) were conducted on the mean number of each type of response, with input, degree of taxonomic relatedness, and gender as between-subjects factors. Table 2 presents the mean number of responses of each type by condition. Given that there were four test trials in each condition, these means varied from 0 to 4. Gender did not have an effect on any of the measures. None of the interactions between factors was significant on any of the measures.

### 2.2.2. Effect of degree of taxonomic relatedness

As predicted, children who were exposed to Strongly related objects were less likely to make mutual exclusivity responses \((M=0.85)\) than children presented with Weakly related objects \((M=1.65)\) \([F(1,32)=4.98, P<.05]\). As expected, the pattern for subordinate responses was the reverse. Children who saw Strongly related objects were more likely to interpret the novel label as subordinate to the familiar label \((M=1.65)\) than children presented with Weakly related objects \((M=0.85)\) \([F(1,32)=5.32, P<.05]\). Degree of taxonomic relatedness had no significant effect on the number of synonymy or familiar-only responses children gave. An analysis of the distribution of children giving mutual exclusivity responses supported the above findings. Only 10 of 20 children exposed to Strongly related object pairs, but 16 of 20 children exposed to Weakly related object pairs, made at least one mutual exclusivity response \([X^2(1, N=40)=3.96, P<.05]\). An analogous analysis on the distribution of subordinate responses was not significant, though the distribution was in the expected direction. Sixteen of 20 children exposed to Strongly related object pairs, and 12 of 20 children exposed to Weakly related object pairs, made at least one subordinate response \((P>.1)\). Finally, there were no item effects. Children made more mutual exclusivity responses for each of the Weakly related pairs than for each of the Strongly related pairs, and conversely made more subordinate responses for each of the Strongly related pairs than for each of the Weakly related pairs.

### Table 2

Mean number of responses (out of four) of each type by condition: Studies 1 and 2

<table>
<thead>
<tr>
<th>Response</th>
<th>Study 1</th>
<th>Study 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inclusive/Strong-rel</td>
<td>Inclusive/Weak-rel</td>
</tr>
<tr>
<td>Mutually exclusive</td>
<td>0.3</td>
<td>1.1</td>
</tr>
<tr>
<td>Subordinate</td>
<td>1.6</td>
<td>1.0</td>
</tr>
<tr>
<td>Synonymy</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>Familiar-only</td>
<td>1.1</td>
<td>0.7</td>
</tr>
</tbody>
</table>
Taken together, these findings reveal that 2-year-olds recognize that objects considered to be very much the same kind of thing taxonomically are more likely to share a name, and are less likely to have mutually exclusive names, than objects considered less likely the same kind of thing taxonomically. Note that not attributing the familiar name to Object B in at least some of the Weakly related pairs can be considered a mistake. That is, a soda bottle, for instance, is in fact a “bottle” even if it is also called “fep.” The finding that children so often made this mistake is evidence for the power of a tendency to expect mutually exclusive names when objects are only weakly related taxonomically.

2.2.3. Effect of input

Input had a significant effect for all but subordinate responses. As can be seen in Table 2, children receiving Inclusive input were less likely to make mutual exclusivity interpretations ($M = 0.70$) than children receiving Exclusive input ($M = 1.80$) [$F(1,32) = 9.19, P < .005$]. Conversely, compared to children in the Exclusive condition, children in the Inclusive condition were more likely to make synonymy responses [$M_{\text{excl}} = 0.10$, $M_{\text{incl}} = 0.70$, $F(1,32) = 7.64$, $P < .01$] and Familiar-only responses [$M_{\text{excl}} = 0.25$, $M_{\text{incl}} = 0.90$, $F(1,32) = 7.70$, $P < .01$].

Again, an analysis of the distribution of children giving mutual exclusivity responses corroborated the above results. Only nine of 20 children in the Inclusive input condition, but 17 of 20 children in the Exclusive input condition, made at least one mutual exclusivity response [$\chi^2(1, N = 40) = 7.03$, $P < .01$]. The difference in terms of subordinate responses was not significant. Fifteen of 20 children in the Inclusive input condition, and 13 of 20 children in the Exclusive input condition, made at least one subordinate response ($P > .4$). We will return to an analysis of how these findings on input relate to the results from other studies in the General Discussion.

In addition to these analyses on specific interpretation types, an ANOVA was conducted using as the dependent measure the mean number of novel labels learned. This learning measure included all items (out of four) for which children picked Object B as the referent of the novel label. Again, input, degree of taxonomic relatedness, and gender were between-subjects factors. None of the factors had an effect on this measure [input: $M_{\text{incl}} = 2.85$, $M_{\text{excl}} = 3.3$, $F(1,32) = 2.37$, $P > .1$; degree of taxonomic relatedness: $M_{\text{strong-rel}} = 2.95$, $M_{\text{weak-rel}} = 3.2$, $F(1,32) = .74$, $P > .3$]. These analyses indicate that differences among conditions on the frequency of specific response types were not due to differences in the overall number of relevant responses.

3. Study 2

In Study 2, we assessed children’s interpretation of labels in the Strongly related and Weakly related object conditions with a neutral input situation. For
both degree of taxonomic relatedness conditions, we provided the children with novel labels but did not relate them to familiar labels as we had in Study 1. We expected a replication of the results of Study 1, with fewer mutual exclusivity responses in the Strongly related than the Weakly related condition.

A second related goal of Study 2 was to provide a more stringent test of 2-year-olds’ sensitivity to the degree of taxonomic relatedness between objects. To that end, we asked children at the end of the session whether an object (Object B in each child’s last set) could be a referent of both the novel label and the familiar basic-level label. This question explicitly addresses whether 2-year-olds accept two labels for an object when they have to consider both labels simultaneously. We hypothesized that 2-year-olds may be willing to accept two labels for an object, when the object is thought to be very much the same kind of thing as an exemplar of a familiar basic-level category.

Finally, a third goal of Study 2 was to provide a baseline for the input conditions of Study 1. Knowing how children interpreted the labels in the absence of any relational input would help us better understand the directionality of the effect of input found in Study 1, in particular, whether it was the case that inclusive input drove children away from clearly mutual exclusivity responses, or whether instead the exclusive input drove them towards such responses.

3.1. Method

3.1.1. Subjects

Twenty 2-year-olds participated in this study ($M = 2$ years 9 months, range = 2 years 3 months to 3 years 1 month). There were 10 girls and 10 boys, from varied socio-economic and ethnic backgrounds. All children were tested in a quiet area of their preschool. Parental consent was obtained prior to testing all children.

3.1.2. Design

Ten children (five boys and five girls) participated in each of two conditions: (a) Strongly related object pairs, and (b) Weakly related object pairs. The mean age of the children in the two conditions did not differ statistically.

3.1.3. Stimuli and procedure

The stimuli consisted of the same eight sets of objects used in Study 1. The procedure was the same as the one used in Study 1 except for two modifications. First, the novel label was introduced without reference to the familiar label (e.g., “This is a dax,” repeated four times while showing a fighter airplane). Second, a question was added at the end of the four trials. Children were asked of Object B (i.e., the referent of the novel label) in the last set whether it could have both the novel and familiar labels (e.g., “Can this be a dax and an airplane?”). As in Study 1, the order of presentation of object pairs was randomized across children; thus, each of the four Object B’s in the two conditions was the target of the
question for at least two subjects. This question was asked only for the last set so as not to bias children’s responses in the earlier trials.

3.2. Results and discussion

3.2.1. Scoring

Children’s responses in the four trials were coded as in Study 1, according to the relation between the novel and familiar labels children maintained (see Table 2 for means). The analyses focused on the two most relevant response types, Mutually exclusive and Subordinate. Children’s responses to the “two labels” question were coded as either “yes” or “no/don’t know.”

3.2.2. Effect of degree of taxonomic relatedness

Separate ANOVAs using degree of taxonomic relatedness and gender as between-subjects factors were conducted on the mean number of mutual exclusivity responses and of subordinate responses. Chi-squares supplemented these analyses.

As in Study 1, children presented with Strongly related object pairs were less likely to make mutual exclusivity responses ($M = 1.9$) than were children exposed to Weakly related pairs ($M = 3.0$) [$F(1,16) = 5.04, P < .05$]. Given the high frequency with which children from both conditions made mutual exclusivity responses, we analyzed children’s individual patterns by looking at the extent to which they made mutual exclusivity interpretations on more than half of the four items.\(^3\) Only four of 10 children in the Strongly related condition made mutual exclusivity interpretations on more than half of the items, whereas nine of 10 children in the Weakly related condition did so (Fisher exact probability = .057). The pattern for subordinate responses was the reverse of the above. Namely, children in the Strongly related condition were more likely to make subordinate responses ($M = 1.1$) than were children in the Weakly related condition ($M = 0.3$) [$F(1,16) = 10.67, P < .005$]. Nine of the 10 children in the Strongly related condition made at least one subordinate interpretation, whereas only three of 10 children in the Weakly related condition did so (Fisher exact probability = .02). The effects of gender and of the interaction between gender and degree of taxonomic relatedness were not significant on either measure. There were no significant effects regarding the other response types.

In addition to these analyses on response types, an analysis was conducted on children’s answers to the “two labels” question. As predicted, children were

---

\(^3\) We initially conducted a chi-square analysis equivalent to the one conducted in Study 1, looking at the distribution of children between conditions who made at least one mutual exclusivity response. This analysis turned out to be uninformative because only one child in each condition of Study 2 did not make at least one mutual exclusivity response. We therefore decided to categorize the children using a more stringent criterion.
more likely to accept that an object could have two labels (i.e., answered the question affirmatively) when the referent was a Strongly related object than when the referent was a Weakly related object. Of the 10 children exposed to Strongly related object pairs, eight said that Object B could be the referent of both a novel label and a familiar label. Only one of 10 children presented with Weakly related pairs responded likewise (Fisher exact probability = .005). In sum, these findings replicate and strengthen the findings from Study 1 regarding 2-year-olds’ sensitivity to the degree of taxonomic relatedness between objects.

3.2.3. Effect of input

Study 2 provided a “Label-only” baseline for analyzing the effect of input on children’s responses. Children’s responses in the present study (N=20) were compared to the responses of children in the Inclusive (N=20) and Exclusive (N=20) input conditions of Study 1. Separate ANOVAs using input and gender as between-subjects factors were conducted on the mean number of mutual exclusivity and subordinate responses from the Label-only condition of Study 2 and Inclusive and Exclusive input conditions of Study 1. Because of the slight age difference of subjects in the two studies, the age in months of the subjects was entered as a covariate in these analyses. There was a significant effect of input on the number of mutual exclusivity responses [F(2,53) = 10.54, P < .001] (see Table 2). Post-hoc Scheffe multiple comparison tests revealed that children in the Inclusive input condition (M = 0.70) made significantly fewer mutual exclusivity responses than did children in the Exclusive input condition (M = 1.80) and in the Label-only condition (M = 2.45). The latter two conditions did not differ significantly (all comparisons at P < .05). As for subordinate responses, input had no significant effect [F(2,53) = 2.23, P > .1] and no two groups differed significantly (Scheffe’s, P > .05). The effects of gender, of the interaction between gender and input, and of the age in months of subjects were not significant on either of these measures.

The finding that Exclusive input did not lead children to make more mutual exclusivity responses than children exposed to the Label-only input contrasts with the results of Merriman (1986). On a closer look, it seems that the main difference between the two studies is that children in Merriman’s Label-only condition made less mutual exclusivity responses than children in the Label-only condition of the present study. One possible reason for this difference has to do with differences in the way children’s name comprehension was tested. In Merriman’s study, children were always tested first on the “broad” label (FTN), and then on the “narrow” label (STN). The present study counterbalanced the testing order of the labels. We decided to do so because of earlier findings suggesting that the order of testing might affect the way children interpret the relation between the labels (Diesendruck, Gelman, & Lebowitz, 1998). Specifically, Diesendruck et al. (1998) found that asking for the familiar/broad label first, as opposed to the novel/narrow label first, decreases the frequency of mutual exclusivity responses. Thus, the difference in the testing procedures of Merriman’s study and the present...
one could account for the lower frequency of mutual exclusivity responses in the former. Taken together, the findings on input in the present studies indicate that 2-year-olds recognize that if “A is a kind of B,” then it is highly unlikely that “A is mutually exclusive with B.”

4. General discussion

The results of the present studies provide convergent evidence about 2-year-olds’ developing understanding of hierarchical relations. This emerging understanding was revealed primarily in two ways: (a) 2-year-olds’ interpretations of the relation between labels varied according to the degree of taxonomic relatedness between their referents; and (b) 2-year-olds recognized that an inclusive input implied a nonexclusive relation between two category labels.

When the referent of a novel label (e.g., a toothbrush of “dax”) and the referent of a familiar label (e.g., a paint brush of “brush”) were not very taxonomically related, children were most likely to interpret the relation between the labels as one of mutual exclusivity. In turn, when the referent of a novel label (e.g., a fighter airplane of “dax”) and the referent of a familiar label (e.g., a passenger airplane of “airplane”) were very much taxonomically related, children’s most common interpretation was of inclusion. These findings corroborate and extend results which showed that whether or not two objects are taxonomically related influences children’s extensions of labels (Tomasello, Mannle, & Werdenchlag, 1988; Waxman & Senghas, 1992).

The findings on degree of taxonomic relatedness may help explain previous results regarding the degree to which 2-year-olds maintain inclusion relations between labels. In particular, it is plausible that subordinate interpretations of novel labels were frequent in Taylor and Gelman’s (1989) study because their novel labels referred to exemplars of subordinate categories of familiar basic-level categories (e.g., a novel label was taught for a green car). In other words, their stimuli resembled the objects in the Strongly related condition of the present studies with regard to their degree of taxonomic relatedness. This similarity between the studies notwithstanding, it is the case that, in contrast to Taylor and Gelman’s findings, the prevalent response in the present study was mutual exclusivity. Two methodological differences between the two studies might help explain this latter disparity. First, Taylor and Gelman used very familiar basic-level categories (e.g., balls and cars), which may have better anchored children’s subdivisions into subordinate categories. Second, the subkinds in Taylor and Gelman’s study were different only with regard to a property (e.g., green versus white cars), thus making a mutually exclusive division less plausible than in the present study.

Degree of taxonomic relatedness also affected children’s willingness to accept two labels for the same object. When a novel label was taught for an object (e.g., a fighter airplane) strongly related to a familiar basic-level category, children
agreed that the object could be called by both the novel ("dax") and the familiar basic-level label ("airplane"). It is possible that in this condition, children realized that the novel label could be denoting a subordinate category. This finding seems to contrast with Blewitt's (1994) observation that 2-year-olds had difficulty with the question of whether an object could have two labels. On a closer look, however, the findings of the two studies are compatible in that Blewitt used objects from different basic-level categories (e.g., chair and table), i.e., objects more similar to the ones used in the Weakly related condition of the present studies. In general, 2-year-olds in the Strongly related condition responded to the two-labels question similarly to the manner in which 3-year-olds (Blewitt, 1994) and 4-year-olds (Au & Glusman, 1990) respond. Taken together, the results regarding degree of taxonomic relatedness demonstrate that 2-year-olds are capable of recognizing a crucial aspect of hierarchical relations. Children understand that objects with a high degree of taxonomic relatedness are likely to share a name, even if one of the objects has an additional name.

The findings on degree of taxonomic relatedness can also indirectly shed some light on how to interpret children’s subordinate responses. A subordinate response consisted of children selecting both target objects (e.g., the sailboat and the motorboat) as referents of the familiar basic-level label ("boat"), and only Object B (e.g., the motorboat) as referent of the novel label ("wug"). One way to interpret this response is as an instantiation of an inclusion relation between the novel and the familiar labels. This type of response, however, is also consistent with other noninclusive overlap relations (e.g., interpreting the labels as denoting categories with common members). The fact that the degree of taxonomic relatedness between the objects affected children’s tendency to make subordinate responses and mutual exclusivity responses hints at the possibility that these subordinate interpretations were indeed reflecting a recognition of the different inclusion relations instantiated in the stimuli. In other words, the finding that these two particular measures were affected by degree of taxonomic relatedness can be taken as indicative that they derived from the same underlying capacity. As noted in the Introduction, the use of mutual exclusivity response variability, in addition to subordinate responding, provides stronger evidence for the argument that 2-year-olds are sensitive to a crucial aspect of hierarchical relations than does the use of subordinate responses alone.

A final but important point to note about the variable of degree of taxonomic relatedness has to do with its derivability. Firstly, children most likely did not infer the status of the objects by simply attending to the physical similarity between objects (which is a reliable cue to degree of taxonomic relatedness; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Adult ratings established that the Strongly related and Weakly related object pairs did not differ in terms of the degree of physical similarity between objects in a pair. Moreover, there were no significant correlations between the adult-established degree of physical similarity between objects, and the frequency of subordinate or mutual exclu-
sivity responses children made to the objects. Secondly, children also could not have relied on the standard linguistic form used to label these objects (e.g., compound versus simple nouns) because some of both Strongly related and Weakly related objects are standardly labeled by compound nouns (e.g., the Strongly related Object B “motorboat,” and the Weakly related Object B “toothbrush”). In fact, this was one of the reasons we chose these particular objects. In sum, the primary dimension on which the two sets of objects differed was in the degree to which adults considered them to be the same kind of thing.

The present studies also revealed that 2-year-olds were sensitive to hierarchical relations instantiated in the relational input used to introduce the novel labels. This conclusion derives not only from the finding that children in the Inclusive input condition of Study 1 were less likely to make clearly mutual exclusivity interpretations than children in the Exclusive input condition, but also from the fact that they were less likely to make such interpretations than children who did not receive any relational input (Study 2). As we reported, age differences did not account for or affect this trend.

On one hand, the finding that 2-year-olds have some grasp of the implication of inclusive input for the hierarchical relation between object labels contrasts with Diesendruck and Shatz’s (1997) results on input. Diesendruck and Shatz found that 2-year-olds were no less likely to interpret two labels as being mutually exclusive when told the novel label was “a kind of” familiar label than when told the novel label was “not a” familiar label. In that study, however, input did influence how many novel labels children learned, suggesting that children discriminated the two types of input. One possible reason for this disparity in the results of the two studies is that in the present work, the naming of exemplars strongly related taxonomically (in the respective condition) might have raised children’s sensitivity to inclusion relations. In Diesendruck and Shatz’ study, most objects were of the Weakly related kind, and thus 2-year-olds did not have the extra cue of degree of taxonomic relatedness to figure out the hierarchical relation between the objects.

On the other hand, the present results are consistent with those of Diesendruck and Shatz (1997) in demonstrating that 2-year-olds do not fully understand the “kind of” phrasing as indicative of an inclusion relation. Although this inclusive input led children to make fewer mutually exclusive interpretations of the labels, it did not significantly increase the frequency of subordinate interpretations — the

---

4 As one reviewer noted, there was a slight — though not significant — difference between Strongly-related object pairs and Weakly related object pairs in terms of adults’ judgments of the degree of perceptual similarity between objects. To test for a possible relation between degree of perceptual similarity and children’s response type, we correlated the degree of perceptual similarity between objects in a pair and the number of mutual exclusivity and subordinate responses raised by that pair, across all children. Although in the expected direction, these correlations were not significant. For mutual exclusivity $r(39) = -0.40, P > .3$; for subordinate $r(39) = .34, P > .4$. Thus, children’s responses were not decisively affected by the degree of perceptual similarity between objects in a pair.
interpretation type that directly follows from such input. Studies have found that by 3 years of age, inclusive input increases the likelihood of children interpreting labels as having a hierarchical relation (Callanan, 1989; Gottfried & Tonks, 1996). Our data suggest a two-step process to achieving this level of interpretive ability, with 2-year-olds able to avoid mutual exclusivity relations in the face of inclusive input but not yet able to impose more regularly a subordinate relation instead.

In conclusion, the present studies suggest that the capacity to recognize hierarchies emerges early. When interpreting the relation between object labels, 2-year-olds have some basic understanding of the relation implied by inclusive statements, and they are mindful of the taxonomic relation between objects. The studies also show that 2-year-olds have the ability to apply multiple labels to objects of a familiar basic-level kind. Nonetheless, inclusive input did not help 2-year-olds as much as it helps 3-year-olds to make inclusive interpretations, and 2-year-olds had more difficulty overall constructing inclusion relations. Two-year-olds may have primarily the capacity to discriminate between hierarchical relations, whereas 3-year-olds already show a capacity to generate such relations, and older children eventually understand the asymmetrical nature of this type of relations (Johnson, Scott, & Mervis, 1997). What may develop between the ages of 2 and 3 years then is the capacity to express an understanding of hierarchies more flexibly and efficiently.

Acknowledgments

This research was partly supported by a fellowship from the Rackham Graduate School of the University of Michigan to the first author. Study 1 is based on part of a doctoral dissertation submitted to the University of Michigan by the first author and was presented as a poster at the Biennial Meeting of the Society for Research in Child Development, Washington, DC, April 1997.

We thank Annemarie Toebosch and Trisha Pizarro for assisting in the collection of the data, the parents and children who came to the University of Michigan Language Laboratory, and the teachers, parents, and children from the following preschools: University of Michigan Children’s Center, in Ann Arbor, MI; Hand in Hand, Little Ranch, and Sandbox, all three in Tucson, AZ.

References


