Factors affecting children’s graphic symbol use in the third year
Language, similarity, and iconicity

Tara C. Callaghan

Department of Psychology, St. Francis Xavier University, P.O. Box 5000, Antigonish, Nova Scotia, Canada B2G 2W5

Received 1 July 1999; accepted 1 July 2000

Abstract

The question addressed in this study is whether the claim that children understand the symbolic status of pictures by the middle of their third year is an overestimate of their ability. Specifically, we asked whether children use language if possible to facilitate their performance in graphic symbolic tasks. Language (availability of verbal labels) was manipulated along with iconicity (degree of resemblance between symbol and referent) and perceptual similarity (between choice items) in a series of four experiments. Children 2.5 and 3 years old were presented with a graphic symbol for 4 s and immediately asked to choose the object depicted (referent) from two choice objects. In Study 1, degree of iconicity between picture and referent was varied and both choice objects had the same verbal label. The 2.5-year-olds failed to use any pictures or replicas as symbols. The 3-year-olds performed well with all types of symbols and better with highly iconic symbols. In Study 2, verbal label availability was manipulated by presenting choice objects having the same or different labels and by varying familiarity of labels. The 2.5-year-olds performed at chance when verbal labels were unavailable but above chance when they were available. The 3-year-olds were above chance in all conditions but performed less well when verbal labels were unavailable. Study 3 confirmed that young children use language to mediate picture symbol use. When 2.5-year-olds were provided with subordinate verbal labels in the matching task, subsequent performance was good even when choice objects had the same basic level verbal label. In Study 4, verbal label availability was contrasted with perceptual similarity between choice objects. When verbal labels could be used and choice objects were dissimilar, performance was best, and when
verbal labels could not be used and choice objects were similar, it was worst. The results suggest that children’s developing understanding of the symbolic function of pictures is tenuous in the third year, and is supported by their use of verbal labels. © 2001 Elsevier Science Inc. All rights reserved.

Keywords: Symbolic development; Graphic symbols; Development of graphic representations

Western culture is awash with pictures that symbolize the physical world. As adults we take the decoding of a picture for granted. This contrasts with words, which are arbitrarily assigned to meaning, and where the recipient must have both the concept and the particular language code to decode the message. These differences may lead one to expect that pictures may be used as symbols before words are; however, there is some indication that the opposite trend holds.

According to DeLoache et al. (DeLoache, 1991, 1995; DeLoache, Pierroutsakos, & Troseth, 1997) and others (e.g., Gibson, 1971), symbolic use of pictures involves the ability to comprehend the essential duality of pictures, the notion that pictures are both 2-D objects in themselves as well as representations of something else. When 2.5-year-old children were shown a picture of a hidden toy and then asked to find the toy in its hiding place in a real room (DeLoache & Burns, 1994), they performed above chance, suggesting they were sensitive to the symbolic status of pictures. When scale models were used instead of pictures, 2.5-year-olds failed to use the symbols but 3-year-olds performed well (DeLoache, Kolstad, & Anderson, 1991). In support of this developmental trajectory, Harris, Kavanaugh, and Dowson (1997) asked children to imagine an outcome using toy and real props (e.g., pig, milk, and ketchup), and to match a pictorial symbol with the imaginary outcome with the props in full view. They reported that 2-year-olds failed, but 2.5-year-olds succeeded, at the task. Taken together, these findings suggest that children understand the symbolic function of pictures by 2.5 years.

Callaghan (1999) argued that these researchers may have overestimated the young child’s abilities, and that understanding of the symbolic function of pictures may not develop until 3 years. A problem in the previous studies (DeLoache, 1991, 1995; DeLoache & Burns, 1994; DeLoache et al., 1991, 1997; Harris et al., 1997), and in symbolic play studies that report children this age use replicas as symbols (e.g., Tomasello, Striano, & Rochat, 1999), is that children were always presented with pictures that depict familiar objects that can be easily labeled. Without a control for the use of language labels by children to facilitate their performance, it is difficult to know whether children are indeed using pictures, or rather are using verbal labels, as symbols to guide performance. Callaghan (1999) asked children to match simple graphic pictures to one of two choice objects, which were deliberately chosen from the same category (i.e., ball) to ensure that children used the picture, and not a word, as a symbol. None of the 2-year-olds, and most of the 3- and 4-year-olds, performed above chance in this
The idea that language, specifically naming, can have an impact on cognitive development has recently gained support from a variety of sources. Children’s knowledge of the name of an object has been found to lead to preferential looking to a picture of that object (Schafer, Plunket, & Harris, 1999), to increased attention to objects (Baldwin & Markman, 1989), and to facilitation of categorization (Waxman & Hall, 1993). During infancy, emergence of words for disappearance is closely tied to success on tasks tapping object permanence understanding, emergence of words for success/failure is related to performance on means/ends tasks, and 18-month-olds’ naming spurt coincides with exhaustive category sorting (Gopnik & Meltzoff, 1997). While there is no direct evidence of how language may be related to the symbolic use of pictures, Wolley and Wellman (1990) report that children’s understanding of the appearance/reality distinction as it applies to pictures — the notion that “it’s not real, it’s just a picture” — does not occur in language samples before 3 years, supporting Callaghan’s (1999) estimate of the onset of graphic symbolism.

One of the difficulties that arises in studies of children’s symbolic processing is designing a task that is indeed symbolic. Perner (1991, pp. 78–82) questions whether DeLoache’s retrieval task requires symbolic understanding of pictures, arguing that children can find the hidden toy on the basis of perceptual correspondence between the picture and the room. However, DeLoache (1991) reports that if pictures are treated more like objects, by cutting them out and hiding a miniature toy under the cutout, children are no longer able to use them as symbols for retrieval. This drop in performance suggests that correspondence alone cannot account for DeLoache’s findings, and supports her contention that children’s ability to conceive of a picture as both an object and a symbol is what develops over the third year. Coming to a mature understanding of the symbolic function of pictures is undoubtedly a lengthy developmental process that involves more than understanding the duality of pictures. Nevertheless, the earliest dawning of graphic symbol understanding may be revealed when the child can use a picture to guide responding in a simple task, and Callaghan’s (1999) research suggests that this emergence may be scaffolded by language. In this series of studies the role of language in graphic symbol use is explicitly examined, along with perceptual similarity between the choice objects in the task and resemblance between a symbol and its referent, in an effort to address concerns with previous research and to extend our understanding of the emergence of graphic symbolism.

1. Study 1: iconicity

DeLoache and her colleagues (DeLoache, 1991, 1995; DeLoache & Burns, 1994; DeLoache et al., 1991) argue that pictorial symbol under-
standing develops sometime around the middle of the third year and that the degree to which a symbol resembles its referent — iconicity — may affect this developing understanding. These researchers assert that before children are symbolic, a high level of iconicity may actually interfere with performance because it is harder for children to suspend consideration of the object status of the picture in favor of the symbolic status, but once they are symbolic, high iconicity should facilitate performance. With 3-year-olds, higher levels of iconicity were found to facilitate the use of 3-D replica symbols (DeLoache et al., 1991), but there have been no direct tests of these predictions using pictorial stimuli. Therefore, in Study 1, 2.5- and 3-year-old children were asked to use a variety of graphic symbols to guide responding in a task where the symbol was removed just before a pair of choice objects was presented. Symbols varied in iconicity. Based on DeLoache’s (1991) model, it was expected that the more the symbol resembled its referent, the harder it would be for 2.5-year-old children to ignore the object status, and attend to the symbolic status, of the symbol. In contrast, it was expected that the performance of 3-year-old children, who do not have a problem with dual representation of pictures, would be facilitated by higher levels of iconicity.

1.1. Method

1.1.1. Participants

A total of 24 children, 12 at 2.5 years (mean = 2 years, 6 months, range = 2 years 4 months to 2 years 7 months) and 12 at 3 years (mean age = 3 years, range = 2 years 9 months to 3 years 5 months), participated in this experiment. There were three boys and nine girls in each of the age groups. Children were recruited through phone contact after consulting birth announcements in a local newspaper and the sample was predominantly white and middle class.

1.1.2. Materials

There were five sets of stimuli used in this study (see Appendix A.1). The Referent set consisted of referent objects (big cats, turtles, tools, shells, rocks, balls, blocks, and Tinker toys) presented for children to choose from in the matching-to-symbol task. The criteria used in choosing and pairing the objects were that they should be relatively familiar and should be readily discriminated within the pairs. Objects were paired (e.g., clam vs. conch shell) so that use of a basic level verbal label (e.g., shell) when making the choice would not facilitate performance. Presumably, to be successful, children would need to use a graphic or amodal representation.

There were four sets of symbolic stimuli that differed in iconicity, or the degree to which they resembled the referent objects. These sets progressively increase the degree of iconicity as illustrated in Fig. 1. The Graphic symbols were simple, black/white, child-like graphic drawings of the objects. Although they
showed minimal graphic equivalence — much as comparing children’s tadpole drawings to real humans — they did not resemble the objects they portrayed in a
realist sense. The *Pencil* symbols were black/white realistic pencil drawings of those objects rendered with perspective and shading. The *Colour* symbols were realistic acrylic paintings of the referent objects painted with perspective by an artist. The *Replica* symbols were 3-D replicas of the objects, all smaller by at least a scale of 1:5 when compared to the referents. To highlight their symbolic status, all pictures and replicas were mounted on white foam board (15.5 cm by 10.5 cm). A toothbrush and spoon, and two graphic drawings of these objects, were used to deliver the instructions for the task.

1.1.3. Design

Each child received all four types of iconic symbols: Graphic, Pencil, Colour, and Replica. There were two stimulus pairs assigned to each of the four conditions for a child, so that each child saw all eight pairs of stimuli, two in each of the symbol conditions. The assignment of stimulus pairs to symbol condition was counterbalanced across children so that all pairs appeared in all possible symbol conditions. Each child received eight trials within a symbol condition, four for each of the stimulus pairs. The stimulus pair chosen for presentation on a given trial was alternated and position of the correct object was randomized. In order to discourage position perseveration, the correct object did not appear on the same side for more than two consecutive trials. Each symbol condition was presented in blocks of eight trials, and order of blocks was counterbalanced across children. In total, each child received 32 choice trials, 8 for each of four symbol conditions.

1.1.4. Procedure

Children were individually tested in a quiet room in a psychology laboratory. The child’s parent was seated nearby and asked not to interact with their child, or provide feedback on his or her performance. A session began with a 10-min familiarization period during which the child played with lab toys and interacted with the experimenter. Following this, the experimenter invited the child to play a game with some pictures. The experimenter first told the child that in this game he/she was to put one thing in the box, which was placed to the side opposite the child’s dominant hand. The toothbrush and spoon were placed on the table top, centered around the child’s midline. The experimenter pointed to one of the objects and asked the child to “put this one in the box.” Then the objects were shuffled about, placed before the child once again and the experimenter pointed to the other object and asked the child to put it in the box. None of the children had difficulty complying with this portion of the instructions. The experimenter then showed the child the pictures of the objects, while the objects were still in view, and asked what was drawn in each of the pictures. All of the children correctly labeled the objects. The experimenter then told the child that “…when we play the game I will show you a picture like this to tell you which thing to put in the box. So if I show you this picture (the experimenter held up one of the pictures), which thing is it telling you to put in the box? And how about this one
(the experimenter held up the other picture)? Which one should go in the box?” The experimenter corrected errors and continued asking the child to put the appropriate object in the box until two consecutive correct responses were made to each of the pictures. The choice task followed immediately.

In the choice task, the experimenter presented the symbol in the center of a screen that shielded the pair of choice objects from view. The screen was placed within arm’s reach of the child in order that pointing and grasping responses could be measured. She held the symbol in place for 4 s, using a clock to time the interval. While holding the symbol, the experimenter asked the child to find the object portrayed in the picture using a variety of prompts, including “Where’s this one?”, “Can you find this one?”, “It’s this one I want you to find”, and “How about this, can you find this one?” When 4 s had elapsed, the experimenter removed the symbol and screen simultaneously, revealing the pair of choice objects. The objects were within reach, and centered about the child’s midline. If the child did not spontaneously do so, he/she was encouraged to put the chosen object in the box. There was only a brief pause between symbol conditions while the experimenter replaced one set of referent objects and their symbols with another. The child was told “I’m going to get some new toys to play the game with,” but were otherwise not alerted to the fact that a new type of symbol would be used. The entire session lasted between 10 and 15 min.

1.1.5. Scoring and analyses
Each child’s mean proportion of correct responses was subjected to t tests to measure order effects and determine whether children were performing above chance, and analyses of variance to assess stimulus counterbalancing and the impact of iconicity on graphic symbol understanding. Post hoc tests of interactions (Tukey’s) in all four studies reached $P < .01$ levels of significance. Corroboration for the ANOVA findings was obtained by linear regression analyses of the data.

1.2. Results and discussion

1.2.1. Order and counterbalancing effects
Because there were no significant effects attributable to order of presentation of the symbolic conditions, this variable was excluded from further analyses. The effects of counterbalancing stimulus pairs to conditions was examined in a $2 \times 4 \times 8$ (Ages $\times$ Assignment $\times$ Stimulus Pair) mixed analysis of variance with Stimulus Pair varied within participants. This analysis revealed main effects for Age (to be discussed below) and Stimulus Pair. Stimulus Pair interacted with Assignment, $F(21,112) = 1.75$, $P < .03$. Only one pair of stimuli was responsible for this interaction; the Tinker Toy pair proved to be more difficult in the Graphic condition than in any of the others. The Tinker Toys were the only pair of stimuli that required a discrimination on the basis of arrangement of parts alone. In retrospect, these stimuli are likely too difficult for children of this age to
discriminate; however, because both ages showed the interaction the data from these pairs were included in the analyses that follow.

1.2.2. Iconicity effects

The young children’s performance was not significantly different from chance, \( t(11) = 0.81, \text{ns} \), which is 0.50 for this two-choice task, while the older children’s performance was significantly above chance, \( t(11) = 14.46, P < .001 \). The degree to which iconicity influenced children’s ability to use graphic symbols was assessed using a 2 x 4 (Ages x Symbol Condition) mixed analysis of variance. The main effects of Age [\( F(1,22) = 2.73, P < .001 \)] and Symbol Condition [\( F(3,66) = 4.92, P < .004 \)] were significant. The Age effect revealed that younger children (0.47) performed less well than older children (0.82) overall. The effect of Symbol Condition showed that the Graphic condition resulted in poorer performance than any of the others, which were equivalent. A linear regression analysis of the data supports the ANOVA results, as indicated in Fig. 2.

Because it was expected that 2.5- and 3-year-old children would be differentially affected by increasing the level of iconicity between symbols and objects, tests of simple main effects for the Age x Symbol Condition interaction [\( F(3,66) = 2.10, P < .10 \)] were conducted (Table 1). The analysis indicated that there was no difference among symbol conditions for 2.5-year-olds, but performance was poor in the Graphic condition and equally good in the other conditions (\( P < .001 \)) for 3-year-olds. Coupled with the \( t \) test results, these findings suggest that 2.5-year-olds did not use any graphic stimuli as symbols

![Fig. 2. Linear regression curves for Study 1 plotted as a function of age and condition.](image-url)
This is inconsistent with DeLoache’s (1991) results for pictures with children this age, but consistent with her results for 3-D models. The findings suggest that 3-year-olds do use graphic stimuli as symbols and support DeLoache’s (1995) claim that pictorial symbols would be more effective with higher levels of iconicity because these children performed less well with the least iconic symbols (simple graphic drawings). There are a number of ways that symbols differed across the conditions of this experiment, including realism, line quality, colour, and form. From Fig. 1 it is evident that the Pencil, Colour, and Replica symbols share the quality of realism with each other and differ in this respect from Graphic symbols. The data suggest that realism is sufficient to provide superior performance and that increasing iconicity beyond that provided in the pencil drawings (Pencil), by adding colour (Colour) or 3-dimensionality (Replica), had no impact on performance.

Although young children did not show differences across conditions when mean proportion of correct responses was analyzed, they did show more frequent touch of the symbols compared with older children (see Table 2). Children at both ages grasped stimuli more often in the Replica condition than in the others ($\chi^2 = 18.7$, $P < .001$ for 2.5-year-olds and $\chi^2 = 18.0$, $P < .001$ for 3-year-olds), concurring with DeLoache’s (1995) and DeLoache et al.’s (1997) claim that young children have difficulty ignoring the object status of symbols the more they resemble objects. It may be that mounting the pictures on foam board in this study highlighted the object status of pictures, much as framed pictures in DeLoache (1991, p. 91) did, and this resulted in failure to use pictures as

<table>
<thead>
<tr>
<th>Symbol condition</th>
<th>2.5 years</th>
<th>3 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Point</td>
<td>Grasp</td>
</tr>
<tr>
<td>Graphic</td>
<td>2.3</td>
<td>0.1</td>
</tr>
<tr>
<td>Pencil</td>
<td>2.9</td>
<td>0.1</td>
</tr>
<tr>
<td>Colour</td>
<td>2.7</td>
<td>0.8</td>
</tr>
<tr>
<td>Replica</td>
<td>0.9</td>
<td>5.3</td>
</tr>
<tr>
<td>Mean</td>
<td>2.2</td>
<td>1.6</td>
</tr>
</tbody>
</table>

Table 1
Mean proportion correct (standard deviation) as a function of age and symbolic condition in Study 1

<table>
<thead>
<tr>
<th>Age</th>
<th>Symbolic condition</th>
<th>Graphic</th>
<th>Pencil</th>
<th>Colour</th>
<th>Replica</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5 years</td>
<td>0.44 (0.20)</td>
<td>0.50 (0.18)</td>
<td>0.45 (0.22)</td>
<td>0.51 (0.20)</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>3 years</td>
<td>0.64 (0.18)</td>
<td>0.85 (0.18)</td>
<td>0.86 (0.12)</td>
<td>0.92 (0.12)</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.54</td>
<td>0.67</td>
<td>0.66</td>
<td>0.72</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2
Mean number (out of a total of eight trials) of points and grasps to symbols as a function of age and symbol condition in Study 1
symbols. However, Callaghan (1999) also reports that 2-year-olds fail to use unmounted pictures as symbols. An alternative explanation of the discrepancy between these findings and DeLoache’s is that 2.5-year-olds use verbal labels to enhance their graphic symbol performance, an opportunity that was deliberately removed in this task but is available in DeLoache’s. This possibility was explored in the following studies.

2. Study 2: language

In Study 2, the availability of verbal labels was manipulated in two ways. First, symbols were presented in three conditions, one that eliminated the opportunity to use verbal labels and two that did not. In the Control condition children could not use verbal labels because members of the choice pair had the same basic level verbal label even though they were perceptually distinct (e.g., cat, cat). In the Standard condition choice objects had different verbal labels (e.g., cat, dog). In this, and the following, condition a strategy of naming pictures could facilitate performance because only one of the choice objects corresponded to the verbal label of the picture. In the Generic condition targets matched the symbols in name but were different exemplars from the objects depicted in those symbols. This condition was included to assess the claim made by DeLoache et al. (DeLoache, 1991; DeLoache et al., 1997) that children may in part fail to understand the dual reality of pictures because of their extensive experience in our culture with the generic referencing function of pictures (i.e., a picture stands for a general category of things, not a particular thing). If this function is salient to children, then they should do well in this condition.

A second way that availability of verbal labels was manipulated in Study 2 was by having choice objects with verbal labels that were either familiar (e.g., cats, rocks, and spoons) or unfamiliar (e.g., octopi, fungi, and corkscrews) to the child. In light of Callaghan’s (1999) findings, it was predicted that 2.5-year-olds would do well only if they could use the verbal label to scaffold performance (i.e., standard and generic conditions, familiar verbal labels) and not when they could not use it (i.e., control condition and unfamiliar verbal labels). It was expected that 3-year-olds would not be as adversely affected by word availability as their younger counterparts.

2.1. Method

2.1.1. Participants

A total of 32 children, 16 at 2.5 years (mean = 30.6 months, range = 29 to 34 months) and 16 at 3.0 years (means = 36.0 months, range = 35 to 39 months) participated in this experiment. There were 7 girls and 9 boys in the younger group and 10 girls and 6 boys in the older group. Children
were recruited as in Study 1, and the sample was predominantly white and middle class.

2.1.2. Materials

Pictures and referent objects were used as stimuli in this study. Pictures were black and white pencil drawings of the choice objects rendered in perspective with shading. Referents were either toy replicas or small everyday and natural objects. Half of the pictures and objects were chosen to be familiar and easily labeled by children this age (dog, cat, spoon, cup, shell, and rock) and half were chosen to be unfamiliar (octopus, lizard, corkscrew, whisk, fungus, and bark). The stimuli included four exemplars (two from each of the subcategories listed above) from each of the animal, kitchen, and natural categories for a total of 24 objects and their corresponding 24 pictures (see Appendix A.2). In order to ensure that perceptual discriminability did not impair performance, choice objects were chosen to be highly discriminable (e.g., whisk and hand beater).

2.1.3. Design

Each child received all conditions in this experiment, and saw both symbols for which they had a verbal label (familiar) and symbols for which they did not (unfamiliar). Thus, both Condition and Familiarity of Verbal Label factors varied within participants. Note that although Familiarity of Verbal Labels for the purposes of analysis was determined for each child at the end of the session, the counterbalancing for this factor was based on the experimenter’s judgment of what was likely to be familiar to children of this age.

In the Control condition, children were presented with a graphic symbol followed by a choice between the object depicted in that symbol and a second object in the same basic level category (e.g., Familiar: German shepherd dog vs. St. Bernard dog, Unfamiliar: birch bark vs. spruce bark). Control trials were blocked together and presented either at the beginning (half the children) or the end (other half of the children) of the sessions. There were six trials for each of the Control-Familiar and Control-Unfamiliar sets of stimuli, one for each of the subcategories of stimuli (e.g., dog). From each pair of stimuli, half of the children received Exemplar 1 as the target (e.g., German shepherd), and the remaining half received Exemplar 2 (e.g., St. Bernard). Half of the children received Familiar stimuli first and the remainder received Unfamiliar stimuli first. Order of stimulus presentation and position of correct object were randomized within Familiar and Unfamiliar blocks.

In the Standard condition, children were presented with a graphic symbol followed by two objects, the one depicted in the picture and the other an exemplar from another category, but within the same global category of animal, kitchen, and natural (e.g., Familiar: German Shepherd dog vs. Orange/White cat, Unfamiliar: Whisk vs. Corkscrew). There were two trials for each subcategory of stimulus in the Standard condition, and both exemplars were used as targets across these trials, giving a total of 12 trials for each of the Standard-Familiar and Standard-Unfamiliar conditions.
In the *Generic* condition, children were presented with the graphic symbol followed by two objects, one that was from the same category but a different exemplar as that portrayed in the picture, and the other an exemplar from another category in the same global category (e.g., Familiar: a picture of a German shepherd was followed by the objects St. Bernard dog and orange/white cat; Unfamiliar: a picture of a whisk was followed by the objects hand beater and corkscrew). There were two trials for each type of stimulus in the Generic condition, and both exemplars were used as targets across these trials, giving a total of 12 trials for each of the Generic-Familiar and Generic-Unfamiliar conditions.

The orders of the two Standard (Standard-Familiar, Standard-Unfamiliar) and two Generic conditions (Generic-Familiar, Generic-Unfamiliar) was determined by a Latin Square, and either preceded or followed by Control trials. Four participants in each of the two age groups received each of the four orders. Order of trials was randomized as was position of the correct object. In order to discourage perseveration, correct objects appeared on the same side for no more than two consecutive trials.

Note that, in the Control condition, it was not possible to mediate symbolic processing with verbal labels, because both items had the same basic level verbal label. It was also not possible to mediate processing with verbal labels for stimuli for which children did not know the words (i.e., all of the Unfamiliar conditions). Additionally, processing of symbols in the Generic-Familiar condition should be helped by verbal mediation because the word was available and was the same for the symbol and referent. Except in the Generic condition, the graphic symbol portrayed an object that was identical to the target objects. (This was not an identity match in the strict sense because the match was between 2- and 3-D stimuli and the symbol was not present during choice.) Insofar as high similarity between symbol and target helps the match, children should perform better in the Standard and Control than in the Generic conditions.

However, similarity between target and foil objects must also be considered, and in this case, low similarity (high discriminability) should help performance. In the Standard condition, choice objects were relatively low in similarity as compared to the Control condition because the exemplars came from different basic level categories in the former and from the same basic level category in the latter. Recall, however, that all stimuli were chosen so as to be highly discriminable when paired for choice and so relative similarity differences were not expected to mediate differential performance across these conditions. To summarize, a number of factors that could affect performance were manipulated in the conditions of this experiment including whether children could (1) access a verbal label for the symbols and targets, (2) make an identity match, and (3) easily detect differences between the choice objects.

2.1.4. Procedure

Testing conditions, instructions to parents, and instructions to children were similar to Study 1. The experimenter presented the picture centered in the
middle of the high side of a stacking box, placed the box within arm’s reach, and centered in front of children so objects could be easily retrieved when the box was turned. The experimenter held the symbol in place for 4 s. In Control and Standard conditions children were instructed to, “Find me the thing that is the very same as this one in my picture.” In Generic conditions children were told that they were to “Find the thing that is kind of like the one in the picture, it’s not the very same but it’s sort of like it.” When 4 s had elapsed, the experimenter turned the box revealing the pair of choice objects. In order to avoid giving direct verbal reinforcement the experimenter simply responded “thank you” when children made their choice but did offer general encouragement that was unrelated to performance throughout the session with comments like “You’re a great finder aren’t you?” or “You’re doing a super finding job!” There was only a brief pause between conditions while the experimenter placed the pictures in random order; however, children were given the opportunity after finishing three blocks of 12 trials (3 conditions) to have a short 5-min break to play with a 3-D bead maze. With this scheduling of pauses and a break all children were able to be fully engaged in the task during test trials.

In order to determine whether children had (Familiar) or did not have (Unfamiliar) verbal labels for the objects used in the task, the experimenter asked children to name objects immediately following the experimental session. If the child failed to give the name of a Familiar object, it was considered Unfamiliar in the analyses (see below), and if they provided the name for what the experimenter considered an Unfamiliar object, this was considered to be Familiar for that child. To be sure that children could also provide the same labels for the pictures as they had for the objects, children were recalled for a short session 2–6 weeks following the completion of the main experiment and asked to name the pictures and objects. The original task lasted between 30 and 40 min, including breaks, and the recall session lasted 5 min.

2.1.5. Scoring and analyses

Each child’s mean proportion of correct responses in each condition (Control, Standard, and Generic) for Familiar Label and Unfamiliar Label stimulus sets was calculated and these means were subjected to analyses of variance to determine whether there were any stimulus or order effects. To assess experimental effects, these data were then adjusted to reflect the actual familiarity of verbal labels for each child. If the child named the pictured and choice objects on a trial it was considered Familiar, and if the child could not label the pictured and choice objects, it was considered Unfamiliar. All children knew most of the labels for the 12 Familiar stimuli (mean = 9.4 for 2-year-olds and 10.5 for 3-year-olds), and very few of the 12 Unfamiliar labels (mean = 1.1 for 2-year-olds and 2.8 for 3-year-olds). All post hoc tests of interactions used Tukey’s method. Linear regression analyses were performed to corroborate ANOVA findings.
2.2. Results

2.2.1. Order and stimulus effects

Order of presentation of the conditions was not significant in the preliminary analysis, and so it was excluded from further analyses. Because there were four trials for each global stimulus category in the Standard and Generic conditions and two trials for each in the Control condition, separate ANOVAs were performed to assess stimulus effects for Standard, Generic, and Control, conditions. In the first ANOVA, the number of correct trials for each of the three types of global categories of objects (Kitchen, Animal, and Natural) used in the experiment was calculated for each child and subjected to a $2 \times 2 \times 2 \times 3$ (Ages x Familiarity x Condition x Category Type) mixed ANOVA, where only Age varied between participants. In the second ANOVA the number of correct trials was analyzed using a $2 \times 2 \times 3$ (Ages x Familiarity x Category Type) mixed ANOVA, with Age varying between participants.

These ANOVAs revealed that Category Type had no impact on performance in the Control condition but did for Standard and Generic conditions. An interaction of Category Type x Familiarity [$F(2,60) = 8.30, P < .001$] indicated that performance with Familiar stimuli was lower for Natural (3.09 out of a total possible 4) than Kitchen (3.53) and Animal (3.48) categories, which were equivalent. For Unfamiliar stimuli, performance was superior for Animal (3.20) as compared to Kitchen (2.52) and Natural (2.75) categories, which were equivalent. In categorization tasks, researchers have reported that North American children perform better with animals than other categories (Mandler, Bauer, & McDonough, 1991), but these results suggest that very familiar kitchen items may also be easily processed. Because Category Type did not interact with any other experimental variable, it was not included in further analyses.

2.2.2. Experimental factors

The mean proportion of verbal labels that children knew for the items that had been assigned to the Familiar set was 0.83 and for the Unfamiliar set was 0.16 averaged across both ages, indicating a small degree of difference between experimenter-defined and child-defined verbal label familiarity. Because it was important to assess children’s graphic symbol processing when the verbal label was accessible and when it was not, each child’s familiar and unfamiliar sets of items was determined, and their mean proportion of correct matches within these sets was calculated for all conditions: Control-Familiar, Control-Unfamiliar, Standard-Familiar, Standard-Unfamiliar, Generic-Familiar, and Generic-Unfamiliar. These data were subjected to t tests to determine whether children were performing at chance levels, and an analysis of variance and linear regression to examine experimental factors. Relevant means are presented in Table 3.
The results replicate and extend the findings of Study 1. Younger children were at chance for Control-Familiar \(t(15) = 1.48, \text{ns}\) and Control-Unfamiliar \(t(15) = 1.61, \text{ns}\) conditions. Older children were above chance with both Control-Familiar \(t(15) = 8.00, P < .01\) and Control-Unfamiliar conditions \(t(15) = 5.73, P < .01\). In the Standard condition all children performed significantly better than chance with both familiar — \(t(15) = 10.02, P < .01\), for younger and \(t(15) = 37.72, P < .01\), for older children — and unfamiliar stimulus sets — \(t(15) = 8.75, P < .01\) for younger and \(t(15) = 12.51, P < .01\) for older children. Young children performed well in the Generic-Familiar condition \(t(15) = 11.78, P < .01\) and although performance was not high, it was above chance for the Generic-Unfamiliar condition \(t(15) = 2.43, P < .05\). Older children were above chance for both Generic-Familiar \(t(15) = 9.39, P < .01\) and Generic-Unfamiliar \(t(15) = 4.55, P < .01\) conditions.

The results of the \(t\) tests suggest different patterns for the younger and older children across the condition and familiarity factors, and in light of these findings, separate analyses for each age group using the same 3 \(\times\) 2 (Condition \(\times\) Familiarity) design were performed. For 2.5-year-olds the ANOVA revealed a significant main effect of Condition \(F(2,30) = 16.81, P < .001\), Familiarity \(F(1,15) = 21.32, P < .001\), and an interaction of Familiarity \(\times\) Condition \(F(2,30) = 4.26, P < .02\). The interaction shows that 2.5-year-olds performed at equivalent levels in the Control condition regardless of familiarity, but performed significantly better \((P < .01)\) in the Standard and Generic compared to Control conditions when stimuli were Familiar. In contrast, the ANOVA for 3-year-olds showed only main effects of Condition \(F(2,30) = 13.39, P < .001\) and Familiarity \(F(1,15) = 16.42, P < .001\). The 3-year-olds were best with Standard and equivalent with Generic and Control regardless of familiarity, and showed significantly better overall performance with Familiar as compared to Unfamiliar stimuli.

Another view of the experimental manipulations can be found in the linear regression results displayed in Fig. 3, which separately plots Familiar and

Table 3
Mean proportion correct (standard deviation) in the choice task of Study 2 as a function of word availability

<table>
<thead>
<tr>
<th>Experimental condition</th>
<th>Control</th>
<th>Standard</th>
<th>Generic</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2.5 years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familiar</td>
<td>0.58 (0.20)</td>
<td>0.84 (0.13)</td>
<td>0.79 (0.10)</td>
<td>0.74</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>0.56 (0.14)</td>
<td>0.73 (0.11)</td>
<td>0.58 (0.14)</td>
<td>0.62</td>
</tr>
<tr>
<td><strong>3.0 years</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Familiar</td>
<td>0.80 (0.15)</td>
<td>0.95 (0.05)</td>
<td>0.81 (0.13)</td>
<td>0.85</td>
</tr>
<tr>
<td>Unfamiliar</td>
<td>0.74 (0.17)</td>
<td>0.85 (0.11)</td>
<td>0.68 (0.16)</td>
<td>0.76</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.67</td>
<td>0.84</td>
<td>0.72</td>
<td></td>
</tr>
</tbody>
</table>
Unfamiliar Label regression data. ANOVA findings were corroborated; there was little improvement over age for the Generic condition, moderate improvement for the Standard condition, and extensive improvement for the Control.
condition. The same trends occurred for both Familiar and Unfamiliar Label stimuli, but the curves are displaced upward for the Familiar as compared to Unfamiliar Label stimuli.

2.3. Discussion

Two findings from Study 2 suggest that language may be relied on significantly more by younger than older children in processing graphic symbols. First, performance is highest for the 2.5-year-olds in Standard-Familiar and Generic-Familiar conditions, both of which allow for verbal mediation. Second, it is lowest when the word is unavailable for mediation, either because it is unfamiliar or because both objects have the same name (Control-Familiar, Control-Unfamiliar, and Generic-Unfamiliar). One finding presents a problem for this interpretation; the relatively high performance for young children in the Standard-Unfamiliar condition. If words were the only mediating factor, then performance should also have been at chance level in this group. There could be two reasons for this effect. The young children may have been able to capitalize here on the relatively low similarity between choice objects and on the identity between symbol and target that is inherent in the Standard conditions. Another possibility is that, although these children did not have the basic level verbal label for these objects, they may have used an attribute word that was enough to help (e.g., for octopi and lizards; the many legged things vs. the long-tailed things), but not enough to have always been reliable and so performance did not reach the high levels attained when a verbal label was easily accessible (i.e., for Familiar stimuli). One problem with this account is that this option is also available for the Generic-Unfamiliar condition where performance was equivalent to the Control conditions. The relatively low performance in the Generic-Unfamiliar compared to Standard-Unfamiliar conditions may be a result of the fact that identity matches between symbol and referent are not possible in Generic conditions.

Older children were relatively unaffected by attempts to derail their performance by making the verbal label inaccessible. Although they performed better in Standard, where verbal labels differed between choice objects, than in Control conditions, they were above chance in all conditions of the experiment. Furthermore, they performed well even if they did not know the verbal label for the stimuli (Unfamiliar), although once again, they were better if they knew the word (Familiar). The chance to make an identity match (Standard) between picture and target object compared to a match between two exemplars of the same category (Generic) did not appear to help performance for the older children. How could 3-year-old children do so well in conditions where the verbal label was inaccessible? They must have been relying on either the accessible image information or they may use subordinate labels to help disambiguate the objects. Older children were helped by favorable similarity relations among the stimuli. They performed best when symbol and target were identical and target and foil objects had different verbal labels (Standard-Familiar). They performed less well
when symbol and target were not identical (Generic as compared to Standard), and when target and foil objects had the same verbal label (Control as compared to Standard). The superior performance across all of the Familiar conditions compared to Unfamiliar ones, along with the better than chance performance in Unfamiliar conditions, argues for the possibility that older children use both verbal and image information when processing graphic symbols.

While the findings of Studies 1 and 2 are robust and suggestive of the use of verbal labels in children’s processing of graphic symbols, they are not conclusive. Nevertheless, some alternative explanations of the main effect can be ruled out. In both Studies 1 and 2, 2.5-year-olds performed poorly in the Control condition. In Study 2 the same children who are performing poorly in the Control condition perform well in the Standard condition. Thus, the poor performance of 2.5-year-olds compared to 3-year-olds is not due to extraneous factors such as fatigue or differential memory abilities in the two age groups. Two possibilities remain. Children did poorly because they could not verbally label pictures in the Control condition, or they did poorly because the choice objects were more difficult to discriminate in the Control compared to the Standard conditions. Study 3 tests the first possibility and Study 4 examines the second.

3. Study 3: subordinate verbal labels

In previous studies, using language to facilitate performance was possible in the Standard condition because children have basic level verbal labels (at least for familiar objects), but was not possible in the Control condition because children of this age range typically do not have the subordinate verbal labels to distinguish the objects. Although some children in Studies 1 and 2 spontaneously labeled pictures when they were presented, it is not possible to tell whether children who did not overtly label pictures were subvocally labeling them. Thus, in Study 3, 2.5-year-olds were given subordinate verbal labels when pictures were presented to further test the hypothesis that young children use verbal labels to scaffold performance in graphic symbol tasks. If performance was improved in the Control condition, then it would converge on the conclusion that children in this age range use verbal labels, if they are available, when processing pictures.

3.1. Method

3.1.1. Participants

A total of 12 children aged 2.5 years (mean = 29.9 months, range = 28 to 32 months) were recruited for Study 3. Two children were excluded from the study — one because of noncooperation and the other because she used both hands to choose both objects on all trials — leaving a total of 10 (7 boys) children. The children were recruited in the same way as previous studies and the sample was predominantly white and middle class.
3.1.2. Materials
As in Study 2, we used black/white perspective drawings and the objects depicted in the drawings. Four objects were chosen from each of two global categories (animal and vehicle). Objects were chosen so that they were familiar, so that subordinate labels could be easily understood and produced by children, and so that they were highly discriminable. The labels and pairings of the objects for Control and Standard conditions are given in Appendix A.

3.1.3. Design
Each child received eight Control and eight Standard trials in counterbalanced order. As in Study 2, Control trials included choice objects from the same basic level category (e.g., tiger cat, cheetah cat) and Standard trials had choice objects from different basic level categories (e.g., tiger cat, sea turtle). Half of the children received Control trials first and the remaining half received Standard trials first. Position of the correct object was randomized but did not appear on the same side for more than two consecutive trials.

3.1.4. Procedure
As in previous studies, children were first given general instructions and then were shown all objects, in pairs, and provided with the verbal label. The experimenter asked the children to look at the objects and, when the objects were attended to by the child, gave the subordinate and basic level labels (e.g., “That’s a tiger cat” or “That’s a cheetah cat”). When all four pairs of objects had been named, the choice task began. Unlike the previous studies, the experimenter gave the verbal label, and asked the child to repeat the verbal label for the pictures on all trials. As the picture was presented, the experimenter ensured that the child was attending to the picture and then said, “This is a (e.g.) tiger cat. Say Tiger cat. (Paused for the child to name the picture, or to decline). OK. Find me the tiger cat.” Then the picture was removed (following 4 s presentation) and the plastic stacking box containing the choice objects was turned to immediately reveal the choice objects. The same labels were used whether Control or Standard trials were presented. There was a short pause between blocks of trials for the two conditions while the experimenter reordered pictures. The entire session lasted approximately 10 min.

3.1.5. Scoring and analyses
Each child’s number of correct choices was determined for each condition and then submitted to correlated t tests to determine (1) whether the level of performance was different from chance for each condition, and (2) whether performance was different across conditions.
3.2. Results and discussion

The mean number of correct choices out of a total possible of 8 was 6.6 (S.D. = 0.70) in the Control and 7.7 (S.D. = 0.67) in the Standard conditions. Performance was above chance levels in both Control \([t(9) = 6.13, P < .01]\) and Standard \([t(9) = 11.76, P < .01]\) conditions. Although levels were high in both conditions, better performance was found for Standard compared to Control \([t(9) = 17.34, P < .01]\).

Taken together with the findings of Studies 1 and 2, these results support the claim that young children use verbal labels when processing graphic symbols. When verbal labels are unavailable, as in the Control conditions of Studies 1 and 2, they perform poorly. When verbal labels are available, as in the Standard conditions of Studies 2 and 3, and the Control condition of Study 3, they perform at a high level. Providing verbal labels can even boost performance in the Standard condition (mean proportion correct is .96 for Study 3 and .84 for familiar stimuli in Study 2), perhaps reflecting that, when unsupported, young children do not always apply strategies consistently. Although the data from any one study alone may not be convincing, the convergence of findings across the three studies is strongly suggestive of the use of verbal labels by children in graphic symbol tasks.

4. Study 4: language and perceptual similarity

A potential confound in the Control and Standard conditions of Studies 2 and 3 is that in addition to a difference in the availability and unavailability of verbal labels, conditions may also differ in the perceptual similarity of objects within the choice pairs. Two cats (Control condition) are more perceptually similar than a cat and a dog (Standard condition). In Studies 2 and 3 this potential confound was addressed by choosing objects that were highly discriminalbe, but the possibility of confound remains. In Study 4, perceptual similarity was manipulated along with availability of verbal labels. Perceptual similarity was manipulated by having choice objects that were easy or hard (but possible) to discriminate and word availability was varied by having choice objects with the same or different verbal labels in Control and Standard conditions. Thus, it was predicted that, as in Studies 2 and 3, accessible verbal labels (Standard) would facilitate performance relative to inaccessible ones (Control). Additionally, the low perceptual similarity of choice objects was predicted to facilitate performance, compared to high similarity. It was also expected that the combination of these two factors, availability of verbal labels and perceptual similarity, may produce a redundancy effect resulting in superior performance in situations where the choice objects were low in perceptual similarity and verbal labels were accessible (Standard-Low Sim), and inferior performance when choice objects were high in similarity and unavailable (Control-High Sim).
4.1. Method

4.1.1. Participants
A total of 32 children, 16 at 2.5 years (mean = 30.5 months, range = 28 to 32 months) and 16 at 3.0 years (mean = 37.0 months, range = 34 to 39 months) participated in this experiment. There were 9 girls and 7 boys in the younger group and 10 girls and 6 boys in the older group. The children were recruited in the same manner as Studies 1 to 3, and the resulting sample was predominantly white and middle class. The children received a small gift as well as the edible stimuli for their participation in this study.

4.1.2. Materials
As in Studies 2 and 3, we used black/white perspective drawings and the objects they depicted. There were three exemplars of each of eight categories of objects (cookie, cracker, pencil, crayon, soap, block, money, and button). These categories were chosen so that they would be familiar to children in this age range, would have verbal labels that were familiar to the children, and could be found in a variety of subtypes, some of which were highly perceptually similar and others that were highly dissimilar so that between category similarity could be varied in the Standard condition. Thus, cookies that looked like crackers, pencils that looked like crayons, soap that looked like blocks, and buttons that looked like money had to be found. The final set included 24 objects, and are described in detail in Appendix A.

Note that for each category one object was considered the anchor and two other objects from that category were chosen as potential contrasts, one that was high in similarity to the anchor and the other that was low in similarity. For example, in the soap category the anchor was Ivory soap, the highly similar object was Dove soap, and the low similarity object was Pears soap. Soaps were only contrasted with other soaps (Control) or with blocks (Standard). The block that was highly similar to Ivory soap was the anchor for the block category. It was painted white, rectangular in shape, and of approximately the same dimensions as the soap. The block that was low in similarity to the soap was a cylindrical red block approximately the same height as the rectangular block.

The edible stimuli in this set were tightly wrapped in plastic wrap so that they could be clearly seen but did not deteriorate with handling. They were replaced for each child.

4.1.3. Design
Each child received all conditions in this experiment. As in the previous studies, in the Control condition children were presented with a graphic symbol followed by a choice between the object depicted in that symbol and a second object from the same basic level category. The second object could be either highly similar (Control-High) to the target (e.g., Ivory vs. Dove...
soap) or low in similarity (Control-Low) when compared to the target (e.g., Ivory vs. Pears soap). In the Standard condition, children were presented with a graphic symbol followed by two objects, the one depicted in the picture and the other an exemplar from another basic level category. The foil could be either highly similar (Standard-High) to the target (e.g., Ivory soap vs. white rectangular block) or low in similarity (Standard-Low) when compared to the target (e.g., Ivory soap vs. Red cylindrical block). The orders of presentation and randomization sequences followed the principles of Study 2. In total, each child received 48 choice trials, 12 for each of four condition × similarity groups.

4.1.4. Procedure

The procedure was identical to Study 2. Children were first given practice to ensure that they understood the instructions; the choice task followed immediately. For both conditions, children were instructed to “Find the one that is the same as this one in my picture.” After the main task, the children were asked to name the objects to ensure that they knew the verbal labels for the objects used in this study. The entire session lasted approximately 25 min.

4.1.5. Scoring and analyses

As in Study 2, each child’s mean proportion of correct choices was determined for each condition. These data were first examined for order and stimulus effects, and then assessed for experimental effects using analyses of variance, which were corroborated with linear regression analysis. Data were not adjusted according to whether children knew the labels (this is addressed in the analyses below). Tukey’s method was used in post hoc analyses.

4.2. Results

4.2.1. Order and stimulus effects

In the preliminary analyses, order of presentation of the conditions was not found to be significant as a main effect or in interaction with other variables. This variable was excluded from further analyses. To assess stimulus effects, the number of correct trials (total possible = 2) for each of the four types of stimulus pairs (Pencil/Crayon, Soap/Block, Cookie/Cracker, and Money/Button) used in the experiment was calculated for each child and subjected to a 2 × 2 × 2 × 4 (Ages × Condition × Similarity × Type) mixed ANOVA, where Age was the only factor varied between participants. A significant interaction of Type × Similarity [F(2,60) = 8.30, P < .001] indicated that, for High Similarity conditions, performance was equivalent across all types, whereas for Low similarity conditions performance was best for Soap/Block contrasts (mean = 1.81 out of a total possible of 2), followed by Cookie/Cracker (1.52), and then by Pencil/Crayon (1.19) and Money/Button (1.17), which
were equivalent. Because Type did not interact with any other experimental variable, it was not included in further analyses.

4.2.2. Experimental factors

The mean proportion of correct matches for the Control-High, Control-Low, Standard-High, and Standard-Low conditions were first submitted to t tests to determine difference from chance. The means relevant to these analyses are presented in Table 4. The 2.5-year-olds were at chance for Control-High \[ t(15) = 1.66, \text{ ns} \] and Standard-High \[ t(15) = 1.51, \text{ ns} \] conditions and significantly above chance for Control-Low \[ t(15) = 2.63, P < .01 \] and Standard-Low conditions \[ t(15) = 7.12, P < .01 \]. The 3-year-olds were at chance for Control-High \[ t(15) = 1.47, \text{ ns} \] and above chance for all other conditions; \[ t(15) = 4.14, P < .01 \] for Control-Low; \[ t(15) = 3.83, P < .01 \] for Standard-High; and \[ t(15) = 8.74, P < .01 \] for Standard-Low.

The \( 2 \times 2 \times 2 \) (Ages \( \times \) Condition \( \times \) Similarity) analysis of variance revealed three main effects and no interactions. The Age effect \[ F(1, 30) = 16.15, P < .001 \] indicated that 2.5-year-olds were performing less well than 3-year-olds (0.55 compared to 0.70). The Condition effect \[ F(1,30) = 33.43, P < .001 \] showed poor performance overall in the Control (0.55) compared to Standard (0.70) conditions. The Similarity effect \[ F(1,30) = 43.60, P < .001 \] was equally robust, with performance at chance for High similarity groups overall (0.53) and above chance when choice items were Low in similarity (0.71). These results are corroborated by the linear regression analysis as presented in Fig. 4.

4.2.3. Verbal knowledge

The stimulus set in Study 4 was unusual because some of the items having different verbal labels were made to look alike. In order to estimate availability of verbal labels, a child was counted as having the word if they labeled any one of the three exemplars with the appropriate label. The 2.5-year-olds knew an

---

**Table 4**

Mean proportion correct (standard deviation) in Study 4 as a function of word availability (Control, Standard) and perceptual similarity (High, Low)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Control</th>
<th>Standard</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2.5 years</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Sim</td>
<td>0.33 (0.12)</td>
<td>0.57 (0.17)</td>
<td>0.45</td>
</tr>
<tr>
<td>Low Sim</td>
<td>0.59 (0.14)</td>
<td>0.72 (0.12)</td>
<td>0.66</td>
</tr>
<tr>
<td><strong>3.0 Years</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Sim</td>
<td>0.57 (0.20)</td>
<td>0.67 (0.18)</td>
<td>0.62</td>
</tr>
<tr>
<td>Low Sim</td>
<td>0.71 (0.20)</td>
<td>0.84 (0.15)</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td>0.55</td>
<td>0.70</td>
<td></td>
</tr>
</tbody>
</table>
average of 6.1 out of 8 words, and the 3-year-olds an average of 6.8 words. The reason that these did not reach the total of 8 words probably had less to do with children not knowing these objects and more to do with the success of manipulating similarity between categories. It was so successful that sometimes children applied the verbal label for one anchor to another one (e.g., called the button anchor money). The 2.5-year-olds showed an average of 2.7 confusions out of a possible 4, and 3-year-olds made an average of 1.9. The low performance at both ages for the Standard-High condition may be related to these confusions. If children labeled both choice objects with the same label due to a confusion, the Standard-High effectively becomes like the Control; there is no word available to scaffold choice. It is not possible to be sure that children were making these confusions at the time they were making their choices because we have no measure of on-line verbal processing. However, the data support the suggestion that potential for disruption of performance is high when verbal labels of choice items are the same.

4.3. Discussion

In Study 4, our manipulation of High similarity was so effective that all but one child confused anchor objects from different global categories at least once. The 2.5-year-olds failed to rise above chance levels even when between-category matches were called for in the Standard-High condition. The 3-year-olds were
also stymied by high similarity, although this was only evident in Control and not in Standard conditions. All children in both conditions benefited from low similarity, or high discriminability, between choice objects.

Children at both ages also showed a strong effect of word availability and no interaction between the two experimental factors. This suggests that both factors, language and perceptual similarity, are operating in graphic symbol processing and both can facilitate or deter performance, depending on the degree to which damage to one processing option is compensated for by scaffolding by the other. If neither the verbal label or a distinctive image are available, then performance suffers as evidenced by chance performance with high similarity pairs in the Control condition at both ages. Three-year-olds were relatively more resilient to attempts to sabotage their performance than were younger children. They were above chance levels in all but Control-High conditions, whereas 2.5-year-olds were above chance only for the Standard-Low condition. What this may reflect is a more flexible processing system for older children, where access to some form of conceptual knowledge is possible (i.e., image or word), as compared to younger children, where knocking out either of the systems places such severe constraints that they are unable to match a referent to its symbol. It appears as though younger children need the redundancy of both language and perceptual systems to help in their choices. Reliance on redundancy may gradually give way to flexible access to alternate forms of conceptual knowledge in graphic symbol processing during the latter half of the third year.

5. General discussion

The findings from this series of studies challenge the accepted view that 2.5 year old children understand the symbolic nature of pictures. Previous research using retrieval (DeLoache et al., 1997) and matching (Harris et al., 1997) tasks failed to control for the possibility that children may use verbal labels to facilitate their performance. The major contribution of the present series of studies is to identify the impact of verbal labels in a picture symbol task by manipulating this variable along with perceptual similarity. The Control task used here should be thought of as a control for language — the child cannot use language in this condition to help pick the object that was represented in the picture because both objects have the same label. The Standard task is what most researchers use to measure understanding of pictures — allow for language to mediate or facilitate performance. The detrimental effect of removing the availability of language in this graphic symbol task (i.e., poor performance in Control relative to Standard conditions) is robust.

Perceptual similarity also has an effect on performance, as illustrated by Study 4. However, the main effect — poor Control relative to Standard performance — occurs even when choice objects are highly discriminable, reflecting a language
mediation rather than perceptual similarity effect. It is possible that both image and verbal representations are called forth when children process graphic symbols of objects in their world. In this research children failed to match symbols with objects when verbal and perceptual information was absent or degraded, and excelled when both forms of information were clear and informative. When only one form of information was available, younger children failed to make the match while older children were apparently able to compensate by using the other form of information. This suggests that flexibility of access to and use of conceptual information develops in the third year.

There is no evidence to suggest that either word or image forms of representations generally dominate graphic symbol processing. Rather, the data indicate that both are called on and that inaccessibility to either will negatively affect performance. There is some indication, however, that, for the youngest children in our studies, using words when presented with graphic symbols may be a favored processing strategy. In all four studies, their performance was usually at chance levels when verbal labels were not available. Very likely, this preference comes from encounters with pictures that resemble the ‘name game’ with words. Specifically, picture interactions between parents and their children usually have the flavor of ‘name the picture’ games, at least early in development, before children are inclined to listen to stories. Perhaps this leads the child to take a naming stance toward pictures. Why are children no longer so reliant on this form of information at the end of the third year? I would argue that by this time their stance toward pictures has changed so that they are now capable of understanding that pictures are sometimes used to symbolize objects (DeLoache, 1991, 1995). To borrow the theory–theory perspective (Gopnik & Meltzoff, 1997), the 2.5-year-olds’ theory of pictures and objects goes something like this: Objects can be named. Pictures can be named. Names can help find things. The 3-year-olds’ theory goes like this: Objects can be named. Pictures can be named, and they can look like objects. When they look like the objects, pictures can help you find them. That 3-year-olds have this more sophisticated understanding is bolstered by the language findings of Wolley and Wellman (1990), but needs to be explicitly tested in studies that measure both graphic symbol performance and appearance/reality understanding in the context of pictures.

How might it be easier for 3-year-olds to make the connection between picture and object? Mandler et al. (Mandler & Bauer, 1988; Mandler et al., 1991; Mandler & McDonough, 1993) has developed a model of early categorization that might be relevant to the choice task used here; a task that is, at its simplest level, a categorization of two stimuli — picture and object. Mandler (1997; 1998) suggests that the essence of an object or event is abstracted from the rich information available to perceptual processes and stored in a simplified analog form called an image schema. Image schemas contain primitive elements of meaning that are used to form the system of
concepts for an individual, but they are not themselves accessible to awareness, although the concepts based on them are. The concepts are accessible in either image or word forms. If this model of infant categorization is correct, and the nature of concept processing is similar in the second and third years of life, then children may apply a categorization process when presented with our choice task. When the symbol is presented, a concept of the depicted object is accessed that is dependent on the child’s theory of pictures — in word form for younger children and in both image and word forms for older children — and then compared with the concepts accessed for the visibly present choice objects. What develops in the third year may be flexible access to both image and word forms of the concept.

DeLoache et al. (1997) propose that symbolic understanding of pictures develops in the third year. The present results suggest that this understanding is tenuous and that early success may be scaffolded by language, and be more apparent than real. The findings suggest that symbolic understanding of pictures emerges late in the third year, and is tied to both language and perceptual factors. Both factors can be harnessed to support graphic symbol use, but if either is unavailable, the tenuous symbolic stance toward pictures fails for the youngest children. Callaghan (1999) reported that social communicative factors also play a role in fostering a symbolic stance toward pictures during the third year. What remains to be explored is how these factors interact later in development to facilitate the development of a more mature understanding of graphic symbols.

Acknowledgments

This research was supported by University Council Research Grant #91973. The author thanks Mary Rankin and Svetoslav Marinov for research assistance and the children for their participation.

Appendix A.

A.1. Stimuli used in Study 1

1. Tools (screwdriver, pliers)
2. Balls (regular, spider — same except black rubber strings emerging from surface)
3. Turtles (plastic sea turtle, wooden decorative)
4. Rocks (granite, sandstone)
5. Big Cats (tiger, cheetah)
6. Shell (conch, clam)
7. Tinker toys (five pieces arranged as stick person, or horizontal bar)
8. Blocks (tall natural cylindrical, green square)

**A.2. Stimuli used in Study 2**

**Familiar**
1. **Animal**
   - Dog (German shepherd, St. Bernard)
   - Cat (orange/white tail up, black tail down)
2. **Kitchen**
   - Cup (plastic juice, metal tumbler)
   - Spoon (metal teaspoon, decorative black/gold serving)
3. **Natural**
   - Shell (moon snail, cockle)
   - Rock (grey/pink granite, White gypsum)

**Unfamiliar**
1. **Animal**
   - Octopus (octopus, squid)
   - Lizard (large collared, small green gecko)
2. **Kitchen**
   - Beater (whisk, hand beater)
   - Corkscrew (red with side levers, silver simple twist)
3. **Natural**
   - Fungus (Fomitopsis pinicola, Grifolia frondosis)
   - Bark (spruce, birch)

**A.3. Sample stimulus arrangements for Study 2**

<table>
<thead>
<tr>
<th>Control-Familiar</th>
<th>Symbol</th>
<th>Choice objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiar</td>
<td>German shepherd</td>
<td>German shepherd</td>
</tr>
<tr>
<td></td>
<td>St. Bernard</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Control-Unfamiliar</th>
<th>Symbol</th>
<th>Choice objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfamiliar</td>
<td>Octopus</td>
<td>Squid</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Standard-Familiar</th>
<th>Symbol</th>
<th>Choice objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiar</td>
<td>German shepherd</td>
<td>Orange/white cat</td>
</tr>
<tr>
<td></td>
<td>St. Bernard</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Standard-Unfamiliar</th>
<th>Symbol</th>
<th>Choice objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfamiliar</td>
<td>Octopus</td>
<td>Collared lizard</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Generic-Familiar</th>
<th>Symbol</th>
<th>Choice objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiar</td>
<td>German shepherd</td>
<td>Orange/white cat</td>
</tr>
<tr>
<td></td>
<td>St. Bernard</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Generic-Unfamiliar</th>
<th>Symbol</th>
<th>Choice objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unfamiliar</td>
<td>Octopus</td>
<td>Collared lizard</td>
</tr>
</tbody>
</table>

**A.4. Stimuli and labels used in Study 3**

<table>
<thead>
<tr>
<th>Vehicles</th>
<th>race car, convertible car</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dump truck, pickup truck</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Animals</th>
<th>sea turtle, land turtle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>tiger cat, cheetah cat</td>
</tr>
</tbody>
</table>

**A.5. Stimuli used in Study 4**

1. **Pencils**
   - Anchor — yellow Dixon 1388-2/HB
   - High Sim — yellow with white birds
   - Low Sim — black with soccer balls and soccer ball eraser
2. Crayons
   Anchor — long, same yellow as pencil anchor, without paper, sharpened
   High Sim — Crayola same colour, with paper
   Low Sim — stubby Crayola, red, with paper
3. Cookies
   Anchor — Peek Frean Lemon Thins
   High Sim — Dare Simple Almond Pleasures
   Low Sim — Peek Frean Fruit Creams
4. Crackers
   Anchor — Ritz original
   High Sim — McCormick’s Country Harvest
   Low Sim — Triscuit original
5. Soaps
   Anchor — Ivory personal size
   High Sim — Dove small
   Low Sim — Pears personal size
6. Blocks
   Anchor — wooden rectangular (4.2 × 8.7 cm) painted white
   High Sim — wooden rectangular (2.2 × 10.2 cm) painted white
   Low Sim — wooden cylindrical column (2.8 cm diameter × 8.8 cm), red
7. Money
   Anchor — Canadian Penny
   High Sim — Italian 1960 three pence
   Low Sim — Canadian Twoonie
8. Buttons
   Anchor — two-holed, metallic, coin-like, same size as penny
   High Sim — two-holed, brass coloured with embossed floral design
   Low Sim — two-holed, plastic red

A.6. Sample stimulus arrangements for Study 4

<table>
<thead>
<tr>
<th>Control-High</th>
<th>Target</th>
<th>Soap anchor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice objects</td>
<td>Soap anchor</td>
<td>Soap High Sim</td>
</tr>
<tr>
<td>Control-Low</td>
<td>Target</td>
<td>Soap anchor</td>
</tr>
<tr>
<td>Choice objects</td>
<td>Soap anchor</td>
<td>Soap Low Sim</td>
</tr>
<tr>
<td>Standard-High</td>
<td>Target</td>
<td>Soap anchor</td>
</tr>
<tr>
<td>Choice objects</td>
<td>Soap anchor</td>
<td>Block anchor</td>
</tr>
<tr>
<td>Standard-Low</td>
<td>Target</td>
<td>Soap anchor</td>
</tr>
<tr>
<td>Choice objects</td>
<td>Soap anchor</td>
<td>Block Low Sim</td>
</tr>
</tbody>
</table>

References


