Journal of Experimental Child Psychology 111 (2012) 230–245

Contents lists available at SciVerse ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Category label effects on Chinese children's inductive inferences: Modulation by perceptual detail and category specificity

Changquan Long [a,b], Xiaoying Lu [c], Li Zhang [a,b], Hong Li [a,b,*], Gedeon O. Deák [d]

[a] Key Laboratory of Cognition and Personality of Ministry of Education, Southwest University, Chongqing 400715, China
[b] Sino-Canadian Research Center for Child Development, Chongqing 400715, China
[c] Zhengzhou Electric Power College, Zhengzhou 450004, China
[d] Department of Cognitive Science, University of California, San Diego, La Jolla, CA 92093, USA

ARTICLE INFO

Article history:
Received 21 September 2010
Revised 15 August 2011
Available online 25 September 2011

Keywords:
Categorization
Cognitive development
Induction
Labeling effects
Perceptual similarity
Stimulus materials

ABSTRACT

Inductive generalization of novel properties to same-category or similar-looking objects was studied in Chinese preschool children. The effects of category labels on generalizations were investigated by comparing basic-level labels, superordinate-level labels, and a control phrase applied to three kinds of stimulus materials: colored photographs (Experiment 1), realistic line drawings (Experiment 2), and cartoon-like line drawings (Experiment 3). No significant labeling effects were found for photos and realistic drawings, but there were significant effects for cartoon-like drawings. Children made mostly (>70%) category-based inferences about photographs whether or not labels were provided (Experiment 1). Children showed a bias toward category-based inferences about realistic drawings (Experiment 2) but did so only when labels were provided. Finally, children made mostly appearance-based generalizations for cartoon-like drawings (Experiment 3). However, labels (basic or superordinate level) reduced appearance-based responses. Labeling effects did not depend on having identical labels; however, identical superordinate labels were more effective than different basic-level labels for the least informative stimuli (i.e., cartoons). Thus, labels sometimes confirm the identity of ambiguous items. This evidence of labeling effects in Mandarin-speaking Chinese children extends previous findings beyond English-speaking children and shows that the effects are not narrowly culture and language specific.

* Corresponding author at: Key Laboratory of Cognition and Personality of Ministry of Education, Southwest University, Chongqing 400715, China. Fax: +86 23 6825 2309.

E-mail address: lihong@swu.edu.cn (H. Li).

0022-0965/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.jecp.2011.08.010

Introduction

Inductive reasoning is a fundamental process in the acquisition of new knowledge and the interpretation of new instances. For example, learning that dogs have *mitochondria* (without knowing what mitochondria are) might lead one to infer that horses also have mitochondria even if one believes that dogs and horses are only distantly related. Because young children have far more limited input and experience than adults and fewer ways of learning about the world (e.g., reading), inductive inference is an especially crucial mechanism for early knowledge acquisition. Thus, understanding how children make inductive inferences is central to understanding cognitive development.

Although there have been many studies of children's inductive inferences, some central questions remain. One of these concerns how children use perceptual similarity (i.e., visual, auditory, and other patterns), as opposed to symbolic or other abstract knowledge, to make inferences. To address this issue, Gelman and Markman (1986) used a classic triad-oddity task. Each trial presented a target item and two comparison items, one of which looked similar to the target and the other of which looked dissimilar but belonged to the same category. Children needed to infer which comparison item shared a nonobvious property with the target. In several studies, 2- to 4-year-olds used category labels instead of perceptual similarity to generalize nonobvious properties (Gelman & Coley, 1990; Gelman & Markman, 1987). This indicated that even very young children use conceptual information to make inductive inferences (Gelman, 2004).

Other investigators have responded that preschool children use conceptual or abstract category knowledge in quite limited ways to support nonobvious inductive inferences (Jones & Smith, 1993; Landau, Smith, & Jones, 1988; Sloutsky, 2003; Sloutsky & Fisher, 2004; Sloutsky & Lo, 1999; Sloutsky, Lo, & Fisher, 2001). Although some studies have replicated the finding that labels facilitate category-based induction (e.g., Graham, Kilbreath, & Welder, 2004; Welder & Graham, 2001), Sloutsky and colleagues (Sloutsky & Fisher, 2004; Sloutsky & Lo, 1999; Sloutsky et al., 2001) argued that children may treat labels simply as perceptual attributes of objects as well as (or instead of) conceptual information. They further argued that labels might be weighted more than visual attributes because they are auditory features, which might be more salient than visual features to infants and young children (e.g., Sloutsky & Napolitano, 2003). Thus, label effects in young children's inductive inferences might actually be based on (auditory) perceptual similarity.

In the current study, we further investigated labeling effects in children's inductive inferences. One goal was to determine whether hearing a label for each of several exemplars causes children to conceptualize the objects as belonging to the named category or whether the label functions as an acoustic similarity cue. We addressed this by comparing same- and different-label conditions. In the different-label condition, unique but familiar basic-level labels were used for each object. For example, one triad had objects labeled "apple," (red) "balloon," and "banana." Here the same-label confound is eliminated because the conceptually related items (i.e., apples and bananas) have different labels. However, in the same-superordinate condition, the apple and banana were labeled "fruit" and the balloon was labeled "toy." Here both conceptual knowledge (about superordinate categories) and label similarity are pitted against visual similarity. By comparing same- and different-label effects, we can separate the role of conceptual knowledge from the perceptual similarity of labels.

To test for the possible effects of labeling in general (e.g., drawing children's attention to the objects or properties), we compared the two labeling conditions with a control condition in which we referred to each stimulus item with the phrase "*zhe4ge4* (这个)", or "this one," combined with a pointing gesture. "*Ge*" is considered a general classifier in Chinese and has few specific semantic implications, although it usually can be applied to countable objects (Myers, 2000). Therefore, the *zhe4ge4* locution controls for verbal attention-getting without providing a discriminative auditory or semantic cue (i.e., it was used for all objects).

How might children's responses to these labeling conditions differ? Several previous studies also used different and same labels on children's induction. For example, Gelman and O'Reilly (1988) used different basic labels (e.g., chair vs. bed) or same superordinate-level labels (e.g., furniture) to test children's induction. In these studies, children were told that one item had a novel property and then were asked to decide whether another thing had the property. The results indicated that children did not reliably generalize properties to labeled superordinate-level items until 7 to 9 years of age (Gelman & O'Reilly, 1988; see also Johnson, Scott, & Mervis, 1997). This suggests that preschool-aged children do not readily use the conceptual relations indicated by superordinate labels. However, Gelman and Markman (1986) used a mixture of basic-level synonyms (e.g., rabbit and bunny) and category-inclusive word pairs (e.g., rose and flower) to test children's inductive inferences. They reported that 4-year-olds made conceptual inductions based on synonymous or category-inclusive (i.e., nonidentical) labels. However, Fisher and colleagues (Fisher, 2010; Fisher, Matlen, & Godwin, 2011) suggested that this might have been due to the fact that some of the label pairs used by Gelman and Markman (1986) were likely to co-occur (e.g., bunny and rabbit), and this could support cued associations rather than conceptual inferences. Indeed, they showed that co-occurrence probabilities can affect children's inductive judgments.

These previous studies do not paint a clear picture of the nature of labeling effects in young children. In particular, it is not clear whether children can use different basic-level labels to induce the conceptual relations that are inherited from a common superordinate category. On the one hand, basic-level labels are learned earlier and verified faster (i.e., are more accessible) than superordinate categories (e.g., Mervis & Rosch, 1981; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Relatedly, preschoolers are less knowledgeable about superordinate words and categories. Smith (1979) found that 4-year-olds sometimes used superordinate categories to make class or property inferences but were less consistent than 5- and 6-year-olds. Other studies reported that preschoolers generalize novel properties to same-basic-level category members more than same-superordinate members (Gelman, 1988; Gelman & O'Reilly, 1988).

Yet even if superordinate labels do not readily evoke conceptual knowledge in young children, they might serve as identical perceptual (acoustic) cues. That is, cue similarity might steer children toward choosing same-superordinate items even if they look different without evoking conceptual (category) knowledge. However, if children *do* use conceptual semantic knowledge to choose same-category/different looking items, different basic-level labels from the same superordinate category should be as effective as, or more effective than, identical superordinate labels. However, if identical labels serve as strong auditory similarity cues (Sloutsky, 2003), different basic-level labels with low co-occurrences should not increase children's conceptual (i.e., taxonomic) inductions, whereas identical superordinate labels should.

These alternative explanations are not, however, mutually exclusive, as Sloutsky and Fisher (2004) suggested. Moreover, labeling effects might not be independent of the stimuli to be judged. For example, labeling effects differ for more and less similar stimuli (e.g., Gelman, 1988; Sloutsky et al., 2001). Other stimulus properties also influence how children make inductive inferences about perceptually similar and/or conceptually related items. For example, Ganea, Pickard, and DeLoache (2008) found that 15- and 18-month-olds more often extended newly learned labels from pictures to objects and from objects to pictures when realistic photos and drawings were presented than when less realistic cartoons were presented. Daehler, Lonardo, and Bukatko (1979) found that 2- and 3-year-olds classified three-dimensional objects better than pictures. They also found an advantage for basic-level over superordinate categories, but they tested only objects, not pictures. Thus, children made inferences differently about objects and pictures, but it was unclear whether this differed for different levels of category relations. Deák and Bauer (1996) also compared 4-year-olds' inductive inferences, with or without identical labels, for objects versus colored drawings of those objects. Children generalized to same-category objects, ignoring perceptual similarity, more with objects than with pictures. This stimulus difference was observed even in the absence of labels. One interpretation is that nonobvious relations are supported by subtle physical properties that are more apparent in objects than in pictures. That is, less detailed stimuli encourage children to generalize based on overall perceptual similarity, whereas more detailed stimuli encourage nonobvious inferences, perhaps due to the richness of physical details. For example, an ornamental seashell-shaped soap looks like a real seashell, especially in a drawing. However, the real object smells perfumed and feels tacky, and these details might indicate its relatedness to a (different-looking) bar of soap even if the to-be-induced properties do not explicitly refer to texture or smell. Deák and Bauer also found that labels promoted conceptual-based inferences about nonobvious properties. Thus, both physically rich stimuli *and* stimulus labels facilitated young children's ability to ''look past'' similar appearances (see Sloutsky & Fisher, 2004).

However, it is unknown how different kinds of labels interact with differently rich and informative stimuli. Cimpian and Markman (2005) suggested, for example, that perceptually simple items are more likely to elicit perceptual-based extension of novel labels. This suggests that labels, particularly different labels, will be less effective at eliciting conceptually based inferences if stimuli are simpler. In less detailed stimuli, features such as shape become more salient. By comparison, detailed stimuli such as objects, or even realistic photographs, provide enough rich perceptual data to confirm labels or to justify non-perceptually-based inferences even if the labels are nonidentical or for superordinate categories. In fact, Deák and Bauer's (1996) results suggest that sufficiently detailed stimuli might obviate labels entirely. However, it is unclear what ''sufficiently detailed'' means. Deák and Bauer's stimuli conflated *detail* with *dimensionality*. That is, the objects were three-dimensional, and three-dimensional items encourage more active exploration. This multimodal exploration process might yield conceptually relevant information more readily even if sufficient information might also be obtained through a more patient (visual) examination of pictures. Thus, it is not clear that stimulus detail per se, as opposed to the process of examining three-dimensional objects, facilitates conceptually based inferences (with or without labels).

This ambiguity raises many intriguing questions about how stimulus properties and perceptual–motor processes affect children's inductive inferences (e.g., Mak & Vera, 1999). We focused on the role of richness of visual information by comparing three kinds of two-dimensional stimuli that have different degrees of visual detail: detailed photographs (Experiment 1), moderately detailed line drawings (Experiment 2), and cartoon-like line drawings (Experiment 3). One question was whether different levels of detail modulate children's tendency to generalize novel properties by conceptual relations versus perceptual similarity. Another question was whether labels affect these tendencies, for which there were two specific questions. First, do labels compensate for lack of visual details by indicating conceptually relevant category relations? Second, do those labels need to be identical (e.g., superordinates such as ''animal'') or do different but related basic-level labels have a similar effect? The answers to these questions will indicate whether preschoolers do indeed use labels to infer abstract category relations.

The current study also addresses a question about language experience. Labeling effects have been studied almost exclusively in English-speaking children. By testing these effects in Chinese children, we explored whether category labels have culture- and language-general effects on young children's inductive inference or whether the effects are population specific.

Experiment 1

We investigated children's inductive inferences about novel properties using colored photos of real objects. In each sets of items, a target looked like one item from a different superordinate category but looked unlike another item from the same superordinate category. There were three label conditions: (a) control phrase (''this one'' was used to name each item), (b) different basic-level labels for each item, and (c) the same superordinate-level labels for the target and dissimilar-looking item. In both labeling conditions, perceptual similarity is pitted against conceptual relations that are implied, directly or indirectly, by the labels.

This experiment has several differences with previous studies. First, some studies (e.g., Gelman & Markman, 1986) did not obtain a baseline like our control condition. Thus, it is not clear how labeling per se affected children's inferences. Although Deák and Bauer (1996) used a similar control phrase, their labeling condition used only identical labels, so it is not clear why the labels were effective. Second, although Gelman and Markman (1986) tested children's responses to nonidentical labels, some of their word pairs had high co-occurrence rates (e.g., rabbit and bunny). This might have affected their data because high co-occurrences can cause each word to cue the other and, by association, the objects (Fisher, 2010; Fisher et al., 2011). To address these concerns, our different-label condition used basic-level labels with low co-occurrence rates. This provides a clean test of the role of conceptual knowledge in children's inductive inferences. Third, to further test how conceptual knowledge versus acoustically identical labels affects children's inferences, we included both a different basic-level label condition and a same superordinate-label condition. If children generalize to the same-superordinate, different-looking items only in the latter condition, it will suggest that they are using identical labels as matching perceptual (acoustic) features rather than as indicators of abstract conceptual relations.

Method

Participants

A total of 161 middle-class children (90 girls and 71 boys) were recruited from a kindergarten on a university campus in Chongqing, China (mean age = 59.2 months, range = 48–72). Of these children, 52 (28 girls and 24 boys, mean age = 59.1 months, range = 48–72) were randomly assigned to the control phrase condition, 56 (30 girls and 26 boys, mean age = 59.6 months, range = 48–72) to the different/basic-label condition, and 53 (32 girls and 21 boys, mean age = 59.0 months, range = 48–72) to the same/superordinate-label condition. Parents provided informed consent in accordance with human participant ethics regulations at Southwest University (Chongqing, China).

In addition, 20 undergraduate students (11 women and 9 men, mean age = 21.5 years, range = 19–25) were recruited from Southwest University. They were paid for participating.

Materials

Six triads of photographs of objects that were familiar to children were used. Each item was printed in color on 15-cm^2 cards. Each triad consisted of a target and two comparison items. One of the latter looked similar to the target but came from a different category; the other looked dissimilar but belonged to the same superordinate category. Note that in some prior studies (e.g., Deák & Bauer, 1996; Gelman & Markman, 1986; Gelman & Markman, 1987), some triads were drawn from the same superordinate category. In these cases, children's responses are hard to interpret because *both* comparison items are conceptually related to the target. To avoid this, in the current study the similar-looking item came from a different superordinate category. For example, in one triad the target, a snake (pinyin: *du2she2*), was categorically related to a different-looking animal, a tortoise (*wu1gui1*) but looked like an inanimate object, a belt (*pi2dai4*). Kindergarten teachers at the children's school were consulted to select the words that would be most familiar to all of the children. Triads are listed in Table 1 along with the Pinyin transliterations of the labels used in the two labeling conditions. Reproductions of the photographs are shown in Appendix A.

The college students rated the perceptual similarity of each target to its comparison items on a scale of 1 (*not similar*) to 7 (*very similar*). The mean similarity between targets and perceptual match items was 5.03 (SD = 0.71); the mean similarity between targets and conceptual match items was 1.65 (SD = 0.39). The difference was significant, $F(1, 19) = 352.50$, $p < .001$, $\eta^2 = .95$. Thus, to adults, perceptual matches looked more similar than conceptual matches. However, because children do not perceive similarity in the same way as adults (Hayes, Heit, & Swendsen, 2010; Sloutsky & Fisher, 2004), 20 children (12 girls and 8 boys, mean age = 51.9 months, range = 47–57) who did not participate in any experiment were asked to choose which comparison item looked more like the target in

Table 1
Items in each triad, by English category name, in Experiments 1 to 3.

	Items: target–perceptual comparison–conceptual comparison	Basic-level labels: target–perceptual comparison–conceptual comparison	Superordinate labels: target perceptual comparison–conceptual comparison
1	snake–belt–tortoise	dushe–pidai–wugui	dongwu–gongju–dongwu
2	apple–balloon–banana	pingguo–qiqiu–xiangjiao	shiwu–wanju–shiwu
3	butterfly–hairclip–ladybug	hudie–faqia–piaochong	dongwu–gongju–dongwu
4	pancake–plate–bread	dangao–panzi–mianbao	shiwu–gongju–shiwu
5	cake–pillow–hamburger	dangao–zhentou–hanbao	shiwu–gongju–shiwu
6	orange–basketball–pineapple	chengzi–lanqiu–bolou	shiwu–wanju–shiwu

Note. Basic- and superordinate-level labels, in Pinyin, are given in the two right-most columns.

each triad. Children chose 86.7% perceptual match items (range = 75–100% across triads). Thus, children perceived the perceptual match items as more similar looking to the targets than the conceptual match items.

To ensure that the item categories were familiar to children, 25 different 4-year-olds (16 girls and 9 boys, mean age = 54.7 months, range = 49–58) were asked to name each item. An experimenter pointed at each photo and asked, "What's this?" Either a basic- or superordinate-level label was considered as correct. Children correctly named 96% of the items (ranges = 88.9–100% across children and 72–100% across items). They mostly (92.6%) produced basic-level labels; the rest were superordinate labels. We then checked children's comprehension of the superordinate labels. Children were shown four items, two from the target superordinate (e.g., *food: apple, banana*) and two noncategory items (e.g., *basketball, belt*), and they were asked to identify the items from a given superordinate. Children averaged 91.0% correct choices (ranges = 82.1–96.4% across children and 80.0–98.3% across superordinates). Thus, children were familiar with the superordinate categories and their labels.

Procedure

Children were tested individually by one of two female experimenters. The unfamiliar properties were "has protein" and "has aminophenol." For each triad, the experimenter told a story about protein or aminophenol: "Scientists found a new substance named [protein/aminophenol]. [Protein/aminophenol] is very little, so only scientists can see it. In this game, you must help the scientist find it. The scientist found that [target] has [protein/aminophenol]. Which of these also may have [protein/aminophenol]?"[1]

The three items then were presented simultaneously, with the target above and the comparison items placed equidistant below. The position of the two comparison items was counterbalanced. The experimenter then pointed to each stimulus and named it (or said "this one" in the control condition) and asked children to repeat each name. This ensured that children had heard the labels correctly.

Results and discussion

Children received 1 point for every conceptual item choice. Conceptual choice scores (i.e., total for each child) were analyzed by analysis of covariance (ANCOVA) with label condition (control phrase, different/basic, or same/superordinate) between participants and age (in months) as a covariate. Conceptual choices averaged 4.17 of 6 (or 69.5%, SD = 22.8%) in the control condition, 4.59 (76.5%, SD = 21.0%) in the different/basic label condition, and 4.68 (78.0%, SD = 22.8%) in the same/superordinate condition. Means (with standard errors) are shown in the three left-most bars of Fig. 1. The labeling condition effect was not significant, $F(2, 157) = 2.14$, $p = .12$, nor was the age covariate, $F(1, 157) = 1.96$, $p = .16$.[2]

Children tended to make conceptual- rather than perceptual-based inductive inferences regardless of labels. Conceptual choices in the control condition were significantly above chance (50%), $t_{\text{two-tailed}}(51) = 6.18$, $p < .001$, $d = 1.17$. This supports Deák and Bauer's (1996) claim that children make abstract category-based inferences even without category labels if they are shown sufficiently informative stimuli (see also Massey & Gelman, 1988). Thus, by 4 or 5 years of age, children can use different kinds of information to make conceptually based inductive inferences even when overall perceptual similarity is in conflict. Given this fact, it is not surprising that the different/basic-level group also was above chance, $t_{\text{two-tailed}}(55) = 9.43$, $p < .001$, $d = = 1.59$. Finally, the same/superordinate condition was above chance (50%), $t_{\text{two-tailed}}(52) = = 8.93$, $p < .001$, $d = 1.68$. These results fail to support the claim that children make inferences based on visual or auditory similarity until 8 to 11 years of age (Sloutsky & Fisher, 2004; Sloutsky et al., 2001).

Other studies have, by contrast, found differences between no-label and same-label conditions (e.g., Jaswal, 2004; Sloutsky et al., 2001). For example, Sloutsky and colleagues (2001) found that preschoolers made more category-consistent inductive inferences in a same-label condition than in

[1] Mandarin text of all instructions and stories is available from the authors on request.
[2] There were no significant triad (item) effects in any group, Wald $\chi^2 = 1.94$, $p = .857$.

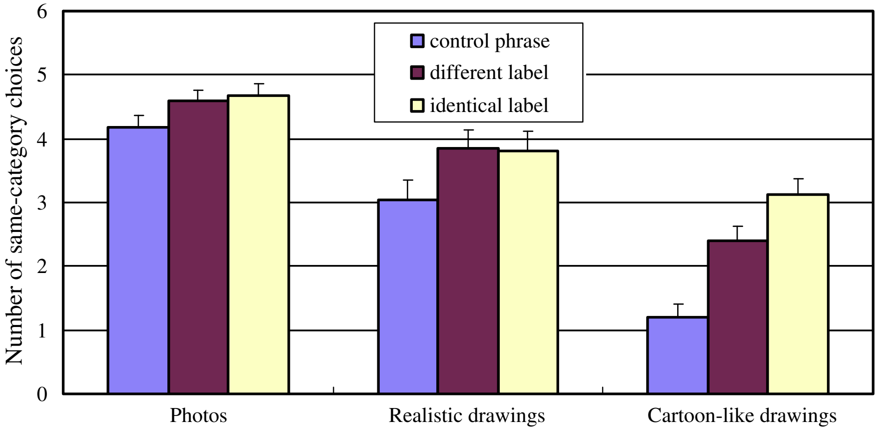

Fig. 1. Children's mean conceptual choices (with standard errors), by labeling condition, in Experiments 1 to 3.

a no-label condition. To reconcile those studies with the current results, we referred to Deák and Bauer's (1996) finding that labels increased children's conceptually based inferences about pictures but were less important for inferences about three-dimensional objects. This suggests that 4-year-olds do not require labels if the stimuli are sufficiently rich and detailed. The color photographs in this study were fairly detailed, and the pretest showed that children could readily identify the objects with the category labels. Apparently, the photographs supported activation of the correct categories, and this rendered the labels unnecessary—perhaps merely confirmatory or redundant. By this token, it is possible that realistic drawings that are moderately detailed, but not as much as photographs, would elicit a larger labeling effect. That is, if there is less physical detail available to verify precise category assignments, labels might have a compensatory role rather than a confirmatory role. Even if this is true, labels might serve that role in different ways, as noted above; identical labels could serve as conceptual symbols, as matching acoustic features, or as both. In any case, superordinates could produce labeling effects. However, if labels activate conceptual knowledge, we might also find labeling effects with different basic-level labels. This would confirm Gelman and Markman's (1986) different-labels results, controlling for a possible confound due to a few high-co-occurrence word pairs.

To test these predictions, we created realistic, but only moderately detailed, black-and-white drawings of the items in the photographs.

Experiment 2

Chinese children's inductive inferences about realistic black-and-white drawings were tested under the same three labeling conditions: different/basic-level labels, same/superordinate labels, and a non-specific control phrase. Deák and Bauer (1996) found that children made fewer conceptually based inferences about two-dimensional drawings than about three-dimensional objects, but their results conflated *detail* with *dimensionality*. This study avoids that problem by using another two-dimensional stimulus. Black-and-white drawings, compared with photographs, carry less category-relevant detail (e.g., texture). By comparing the results with those of Experiment 1, we can test whether labels compensate for category-relevant stimulus details in sensitizing children to abstract conceptual relations.

Method

Participants

A total of 93 children (mean age = 56.4 months) who had not previously participated were recruited from the same kindergarten. Parental consent was obtained. Of these children, 35 (17 girls and 18 boys, mean age = 57.8 months, range = 52–68) were randomly assigned to the control condition, 32 (17 girls and 15 boys, mean age = 56.8 months, range = 48–71) to the different/basic condition, and 26 (17 girls and 11 boys, mean age = 56.6 months, range = 48–71) to the same/superordinate condition.

Materials and procedure

An artist made black-and-white charcoal drawings of the photographs used in Experiment 1. These are shown in Appendix B. Another 20 children (14 girls and 6 boys, mean age = 52.6 months, range = 48–58) judged which comparison item looked most like the target, as in Experiment 1. Children chose 83.3% perceptual matches (range = 75–95% across triads). Thus, children perceived the perceptual matches as more similar looking. The procedure was otherwise identical to than in Experiment 1.

Results and discussion

Children received 1 point for every conceptual match choice. Scores were analyzed by ANCOVA with label condition between participants and age as a covariate. Children averaged 3.06 of 6 (or 51.0%, *SD* = 30.6%) in the control condition, 3.84 (64.1%, *SD* = 27.8%) in the different/basic condition, and 3.81 (63.5%, *SD* = 26.7%) in the same/superordinate condition. These are shown in Fig. 1 (middle bars). The condition effect was not significant, $F(2, 89) = 2.13$, $p = .13$. Thus, as in Experiment 1, there was not a reliable labeling effect.[3] The age covariate was marginally significant, $F(1, 89) = 3.48$, $p = .07$.

Despite the nonsignificant group effect, children in both the different/basic-level and same/superordinate conditions made more conceptual choices than chance, $t_{\text{two-tailed}}(31) = 2.86$, $p < .01$, $d = 0.84$, and $t_{\text{two-tailed}}(25) = 2.57$, $p < .05$, $d = 0.81$, respectively, whereas children in the control condition did not, $t_{\text{two-tailed}}(34) < 1.0$, $p = .85$. This finding, along with that of Experiment 1, indicates that Mandarin-speaking children, like English-speaking children in previous studies, can make conceptually based generalizations of novel properties, even with nonmatching labels. However, because absolute levels of conceptual generalizations differ fairly broadly across studies, we cannot precisely compare the current results with those of previous studies with English-speaking children. Regardless, taken together with Experiment 1, the results imply that labels do not necessarily cue conceptual relations; they produced only a trend (and an above-chance bias) with drawings.

From these results, it remains unclear what role labels play. One possibility was suggested by Sloutsky and Fisher (2004) and Jaswal (2004). Those studies used detailed realistic images that were progressively morphed between two basic-level categories (e.g., cat and dog). The differences between exemplars were quite subtle. Both studies found a significant labeling effect. Gelman and Markman (1987) also found labeling effects—but with stylized or "cartoonish" line drawings that might have been hard to identify without labels (see also Farrar, Raney, & Boyer, 1992; Florian, 1994). Although the morphed photograph stimuli are quite different from stylized cartoonish stimuli, a common factor might be that the relevant categories are hard to identify—in the former because between-exemplar differences were subtle, and in the latter because there was not enough category-specific visual information. This suggests that labeling effects are not specific to dimensionality (i.e., two- or three-dimensional) or to a specific kind of label (e.g., basic or superordinate level). Labels are effective *if* they help to classify hard-to-identify items, where "hard to identify" depends not on the stimulus type per se but rather on the stimulus detail *relative to* the difficulty of the discrimination or identification task. If pictures do not support clear classification, labels might take on a compensatory role. Put differently, as stimuli become less informative (or more ambiguous), labels can become more useful. To test this hypothesis, in Experiment 3 children made inferences about cartoon-like drawings. Cartoons can differ from photographs and realistic drawings in two ways: they lack detail and they are stylized, so that category-diagnostic features may be distorted.

Experiment 3

In this experiment, cartoon-like line drawings were used to test children's inductive inferences. Children heard a control phrase, different/basic labels, or same/superordinate labels, as in the Exper-

[3] There were no significant item (triad) effects in any group, Wald $\chi^2 = 6.18$, $p = .289$.

iments 1 and 2. The drawings were stylized representations of the same categories. They were drawn to be similar to the pictures of animals and objects used in some children's cartoons. This, incidentally, raises a separate issue—abstract cartoons occur not only in television and print cartoons but also in educational books and media, where stylized images are used to draw attention to certain stimulus properties. Therefore, simplified cartoons are a prominent medium for children's education as well as entertainment. This is a practical rationale for testing children's reasoning about cartoons as "socioecologically" meaningful stimuli.

The cartoon drawings were described using the same labels as Experiments 1 and 2. They were drawn with a few details to support abstract conceptual classification (e.g., eyes, suggesting an animal type), and these features always distinguished the target and conceptual match from the perceptual match. Thus, it was feasible to infer that the different-looking pictures shared a nonobvious property, although such details seemed subtle compared with overall visual similarity.

Method

Participants

A total of 164 children (mean age = 59.7 months) who did not participate in either Experiment 1 or 2 were recruited from the same kindergarten with parental consent. Of these children, 56 were randomly assigned to the control phrase condition (31 girls and 25 boys, mean age = 60.2 months, range = 48–71), 54 (30 girls and 24 boys, mean age = 59.2 months, range = 48–71) to the different/basic-label condition, and 54 (30 girls and 24 boys, mean age = 59.8 months, range = 48–70) to the same-label condition.

Materials and procedure

Cartoon-like drawings (Appendix C) were displayed on cards of the same size as the photographs. They were drawn to be abstract versions of the categories used in Experiment 1. Drawings were modified to yield overall perceptually similarity ratings comparable to the ratings of the photographs. The same college students who made similarity judgments in Experiment 1 also rated the perceptual similarity of each target drawing to both comparison items on a scale of 1 (*not similar*) to 7 (*very similar*). The overall mean target to perceptual match similarity was 5.17 ($SD = 0.81$), and the mean target to conceptual match similarity was 1.64 ($SD = 0.41$). This difference was significant, $F(1, 19) = 601.6.$, $p < .001$, $\eta^2 = .97$. Also, similarity ratings for photos and cartoon-like drawings were not significantly different, $F_{\text{perceptual item}}(1, 19) < 1$, $p = .51$, and $F_{\text{conceptual item}}(1, 19) < 1$, $p = .94$, respectively. Thus, the relative perceptual similarity within the cartoon-like triads was similar to that of the photographs.

To determine whether children knew or could identify the stimulus categories, the same preschoolers who did the familiarity pretest for photos in Experiment 1 also did a familiarity test for the cartoon-like drawings. For each drawing, children were first asked to name it and then asked whether it belonged to a superordinate category (by choosing it from an array of four drawings) and a basic-level category. Children often gave unrelated labels or said "I don't know," averaging 33.6% correct ($SD = 19.6\%$) for superordinate categories and 35.8% correct ($SD = 26.9\%$) for basic-level categories. Correct spontaneous labels ranged from 11.9% to 68.3%. Thus, in contrast to the photos in Experiment 1, children could not always identify the intended category of each cartoon drawing. This confirms that the cartoons depicted less category-relevant information. Consequently, Experiment 3 should be a strong test of labeling effects.

Results and discussion

Children received 1 point for every conceptual choice. Their totals averaged 1.20 of 6 (or 20%, $SD = 26.7\%$) in the control condition, 2.41 (40.2%, $SD = 26.0\%$) in the different/basic-level condition, and 3.13 (52.2%, $SD = 30.8\%$) in the same/superordinate condition. These means are shown in the right-most bars of Fig. 1. Scores were analyzed by ANCOVA with condition between participants and age as a covariate. The age covariate was not significant, $F(1, 160) < 1$, $p = .37$. However, there was a significant label condition effect, $F(2, 160) = 18.55$, $p < .001$, $\eta^2 = .19$, which was similar to previous studies (e.g., Davidson & Gelman, 1990; Deák & Bauer, 1996; Farrar et al., 1992). Posterior pairwise comparisons (LSD) showed significant differences between each pair; children made more conceptual choices in both labeling conditions than in the control condition, and they made more in the same/superordinate condition than in the different/basic condition.[4] This confirms that identical labels can partly override perceptual similarity (e.g., Farrar et al., 1992; Sloutsky et al., 2001).

Conceptual choices were compared with chance (50%). Children in the control condition chose fewer conceptual items (i.e., more perceptual matches) than chance, $t_{\text{two-tailed}}(55) = -8.43$, $p < .001$, $d = -1.80$. Thus, as predicted, when category information was degraded and no labels were provided, children generalized properties based on perceptual similarity. In the different/basic-label condition, children also were below chance, $t_{\text{two-tailed}}(53) = -2.79$, $p < .01$, $d = 0.59$. Because children understand the basic-level words (Experiment 1 pretest), this suggests that the labels did not convince children of the items' category identity. In the same/superordinate-label condition, however, children did not generally respond based on perceptual similarity; rather, they were not different from chance, $t_{\text{two-tailed}}(53) = 0.51$, $p = .61$. This suggests that identical labels eliminated, but did not reverse, children's tendency to generalize by appearances. Several previous studies also reported that identical labels do not compel children to make mostly conceptually based inferences about hard-to-identify or novel items (e.g., Davidson & Gelman, 1990; Farrar et al., 1992). These results have an important implication for the design of educational materials. Stylized drawings are often used in educational texts or videos for children, presumably because they are attractive or entertaining. However, such images might sometimes bias children to discount labels or descriptions of category information. If our finding is general, it will suggest that factual information about real categories is best supported by realistic and detailed images that offer *some* perceptual support, even if subtle, for a category-based generalization. Cartoons can simplify and amplify critical features, but they might sometimes impede children's acceptance of category-relevant information.

These results can mediate a debate about the role of category labels. Sloutsky, Kloos, and Fisher (2007) also found that even if children *can* categorize items in adult-like ways (as in Experiments 1 and 2), they mainly use perceptual similarity to guide property induction. The authors used artificial animals and created a rule so that category membership was not correlated with appearance (e.g., "Ziblets" have more fingers than buttons, "Flurps" have more buttons than fingers). After children learned this rule, they completed an induction task. Children tended to respond based on perceptual similarity rather than on category membership. However, Gelman and Waxman (2007) argued that Sloutsky and colleagues' (2007) categories were artificial and so children might not have treated them as typical categories. The current study avoided that problem by using real familiar categories. Nevertheless, the results are partly consistent with Sloutsky and colleagues' argument—an identical acoustic feature (i.e., label) had a stronger effect than different/basic labels, $t_{\text{two-tailed}}(103) = 2.21$, $p = .030$. Because children produced the basic-level words more readily than the superordinates (Experiment 1 pretest), the finding suggests that identical labels carry greater salience.

However, the different/basic labels also significantly reduced the perceptual bias compared with the control condition, $t_{\text{two-tailed}}(108) = 4.06$, $p < .0001$. This is partly consistent with the claim that children treat labels as conceptual information (Gelman & Waxman, 2007). Because the labels were not confounded by high co-occurrence rates, it seems that *conceptual* information conveyed by labels occasionally pushed 4-year-olds to ignore perceptual similarity even for hard-to-identify stimuli. However, the labels did not overall reverse appearance-based responding or increase the number of children who reliably responded based on labels; in the control and different/basic conditions, only 7.1% and 7.4% of children, respectively, made 5 or 6 conceptually based inferences versus 29.6% in the superordinate condition. Sloutsky and colleagues (2007) similarly found that labels significantly reduced a perceptual bias compared with a control condition. Thus, there is converging evidence that labels can, depending on other factors (e.g., perceptual similarity, category familiarity), either compel 4-year-olds to override perceptual similarity or simply weaken their use of similarity. In sum, labels were not used as a top-down basis for taxonomic induction; rather, they were used as one of many probabilistic cues.

[4] There were no significant between-triad differences in any group, Wald $\chi^2 = 6.11$, $p = .296$.

General discussion

Preschoolers' inferences about nonobvious properties of photos, realistic drawings, and cartoon-like drawings were measured in a triad oddity task that pitted perceptual similarity against category membership. Effects of basic-level and superordinate labels were tested. The labeling effect was reliable only for cartoon-like drawings. For photographs, children extended properties according to conceptual relations whether or not the items were labeled. For realistic drawings, children made a majority of conceptual inferences about labeled items but were at chance with unlabeled items. For cartoon-like drawings, labels reduced or eliminated a tendency to respond based on similarity of appearance.

These results confirm that labeling effects depended on the nature of the stimulus (Deák & Bauer, 1996; Ganea et al., 2008). The results also show that the dependency is not based strictly on dimensionality (because all materials were two-dimensional) but also on the richness of stimulus detail. Photographs, the most detailed stimuli, obviated labels. This confirms Deák and Bauer's (1996) argument that labels are not the only factor that can compel children to generalize in terms of abstract categories. Subtle stimulus details can have a similar effect. It seems that labels are most effective when stimuli are hard to identify with respect to a particular judgment. For example, the labeling effects reported by Davidson and Gelman (1990), Deák and Bauer (1996), Farrar and colleagues (1992), and Gelman and Markman (1987) all involved simple cartoonish line drawings. However, Jaswal (2004) and Sloutsky and colleagues (Sloutsky & Fisher, 2004; Sloutsky et al., 2001) found labeling effects for photographic stimuli. This fits the argument that labels interact with stimulus detail relative to the difficulty of the particular judgment. Jaswal and Sloutsky and colleagues' participants needed to make nuanced judgments about morphed intermediate items that were subtly different. In these cases, labels can reinforce judgments about subtle stimulus details. Thus, those studies and the current one suggest that labeling effects depend on whether the stimuli provide enough information to clearly make an identification to justify a conceptual inference.

This explanation helps us to adjudicate between two alternative views of children's use of perceptual similarity versus conceptual relations. One view is that children typically generalize based on perceptual similarity. The other is that children are guided by specialized concept-relevant information such as labels. With respect to the latter, studies of novel word extension have shown that children relate category labels to some stimulus dimensions more than to others (e.g., Imai, Gentner, & Uchida, 1994; Landau et al., 1988). Sloutsky and colleagues (2001) suggested that labeling effects could be due to children treating labels as matching auditory features. However, Gelman and Coley (1990) argued that this is implausible because Gelman and Markman (1986) showed that synonymous labels also have an effect and because children sometimes make category-based inferences without labels. Fisher and colleagues (Fisher, 2010; Fisher et al., 2011) argued that the synonymous items in Gelman and Markman (1986) were high co-occurrences pairs (e.g., rabbit and bunny) that could have skewed the results. Gelman and Coley (1990) did not fully address the issue. In the current study, we avoided this by using a different basic-level condition. The results conditionally confirm Gelman and Coley's claims by showing that with photos and realistic line drawings, children can make conceptual choices even when given different basic-level labels.

The results also extend previous work by showing that Mandarin-speaking children can make category-based induction without labels. Saalbach and Imai (2006) also tested Mandarin-speaking children's novel property generalizations using color drawings. Their comparison items included similar-appearance and same-superordinate items, like ours, but they also included a third thematic comparison item. They did not use labels. Their results are consistent with ours in Experiment 1 (photographs) and with those of Deák and Bauer (1995), Deák and Bauer (1996). Moreover, several studies found that the shape bias—common in object–word extension tasks—could be eliminated in some cases (e.g., Booth, Waxman, & Huang, 2005; Cimpian & Markman, 2005). For example, Cimpian and Markman (2005) reported that the shape bias can be reduced by adding another comparison item or by increasing the complexity of the objects. These results generally confirm and extend Gelman and Coley's (1990) arguments.

However, we also found that with stylized drawings, children did not make mostly conceptual inductions even with familiar labels. The finding that category labels did not always compel children to ignore appearances might appear to be inconsistent with previous findings, but this is not necessarily the case. For example, Gelman and Coley (1990) found that familiar labels pushed 2-year-olds to make more conceptual-based inductions about atypical (i.e., hard-to-identify) items. However, their category *and* property labels were familiar, and at least some of those paired labels likely had high co-occurrences or semantic associations; five of the nine pairs were rabbit/carrot, dog/barks, bird/nest, leaf/tree, and flower/garden. These words likely primed one another and might have elicited more "yes" responses even for different-looking drawings. Moreover, participants were shown drawings one by one, so they were never directly confronted with a conflict between different-category/similar-looking and same-category/different-looking items. It is not clear whether Gelman and Coley's labeling effects would generalize to a task with novel property words and concurrent stimulus presentation.

Despite these caveats, the results of Gelman and Coley (1990) and several other previous studies, taken with the current results, sketch a plausible developmental account—labels act as a kind of *conditional* "cumulative cue." If labels are the *only* evidence favoring a conceptual relation, and there is other conflicting evidence or simply insufficient support for the labels, children might ignore them. For example, Davidson and Gelman (1990) reported that preschoolers did not use novel labels to generalize properties of the unfamiliar items if the labels were too discrepant with children's classification of the items in the absence of labels. Similarly, our stylized drawings in Experiment 3 were difficult to identify, and this might have caused children to ignore or discount the labels. In fact, children do not always believe and accept adults' labels (e.g., if an adult calls a dog "cat") (Gelman, 2009; Jaswal, 2004). Thus, even young children weigh the plausibility of labels in light of other information.

The current results also address a debate about whether category labels serve as acoustic similarity features, as symbols for category membership, or as both. The results suggest that they serve both roles in context-specific ways. In Experiment 3, different basic-level labels elicited more conceptual responses than baseline but did not reverse children's reliance on appearances. However, identical superordinate labels elicited a greater number of conceptually based inferences (but still did not have a dominating effect). This suggests that children are more strongly compelled by labels that are acoustically similar. Thus, labels have a "dual nature," functioning in ambiguous situations as acoustic cues and/or as signals of category membership.

A final implication of our results is that previous studies of inductive inferences with English-speaking children are consistent with results from Chinese children speaking the Chongqing dialect of Mandarin Chinese (重庆话). Mandarin has different lexical and semantic properties than English (e.g., classifiers, extensive homophony), so it could not be assumed that labeling effects found in English-speaking children would generalize. Although Saalbach and Imai (2006) found an age-related shift in Chinese preschoolers' use of perceptual, thematic, and taxonomic information for inductive inferences, that study did not test labeling effects on novel property inferences. However, their results and ours extend previous results to children learning a non-Indo-European language. Therefore, it is plausible that some labeling and stimulus effects on children's inductive inferences are language general.

In sum, preschoolers made predominantly conceptually based inferences about relatively detailed stimuli such as color photographs. Effects of both different basic-level labels and superordinate-level labels were limited to items that were hard to identify; specifically cartoon-like drawings. Thus, perceptual similarity and labels combine in nonlinear ways to guide young children's property generalization. The results support the view that there is no qualitative shift from perceptual to conceptual induction (Deák & Bauer, 1996). Rather, children's responses are systematically adapted to their existing knowledge, available perceptual information (visual and auditory), and plausible linguistic cues.

Acknowledgments

This research was supported by the Fundamental Research Funds for the Central Universities (SWU1009091, NSFC30900408, and NSKD08003). Thanks go to Mutsumi Imai and the reviewers for insightful comments on an earlier version of the manuscript. We thank the children and teachers of Southwest University kindergarten for their participation. We also thank Ma Xiaoqing, Ouyang Hanlu, Wu Yin, and Feng Tingyong for their assistance.

Appendix A

See Fig. 2.

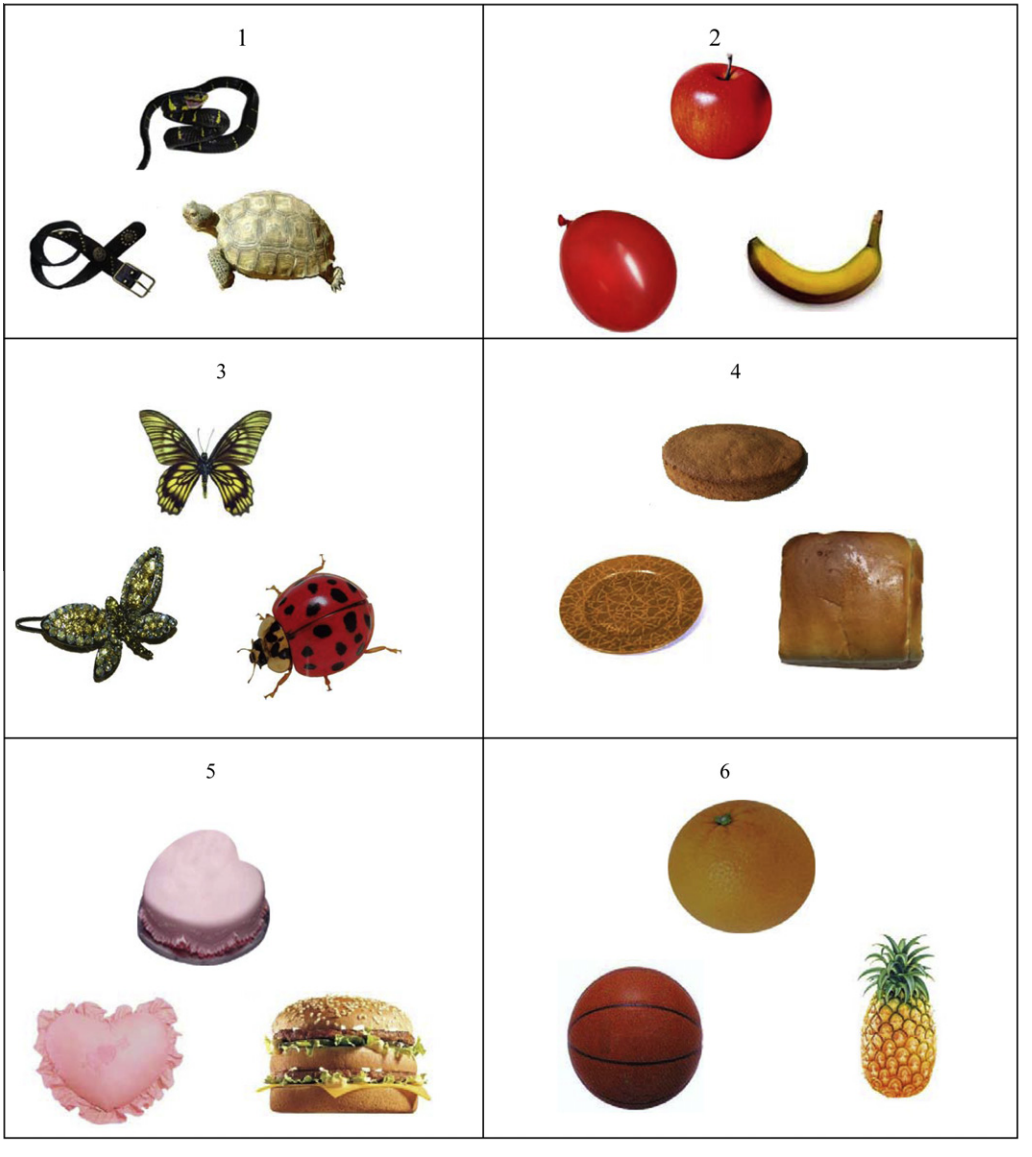

Fig. 2. Triads in Experiment 1.

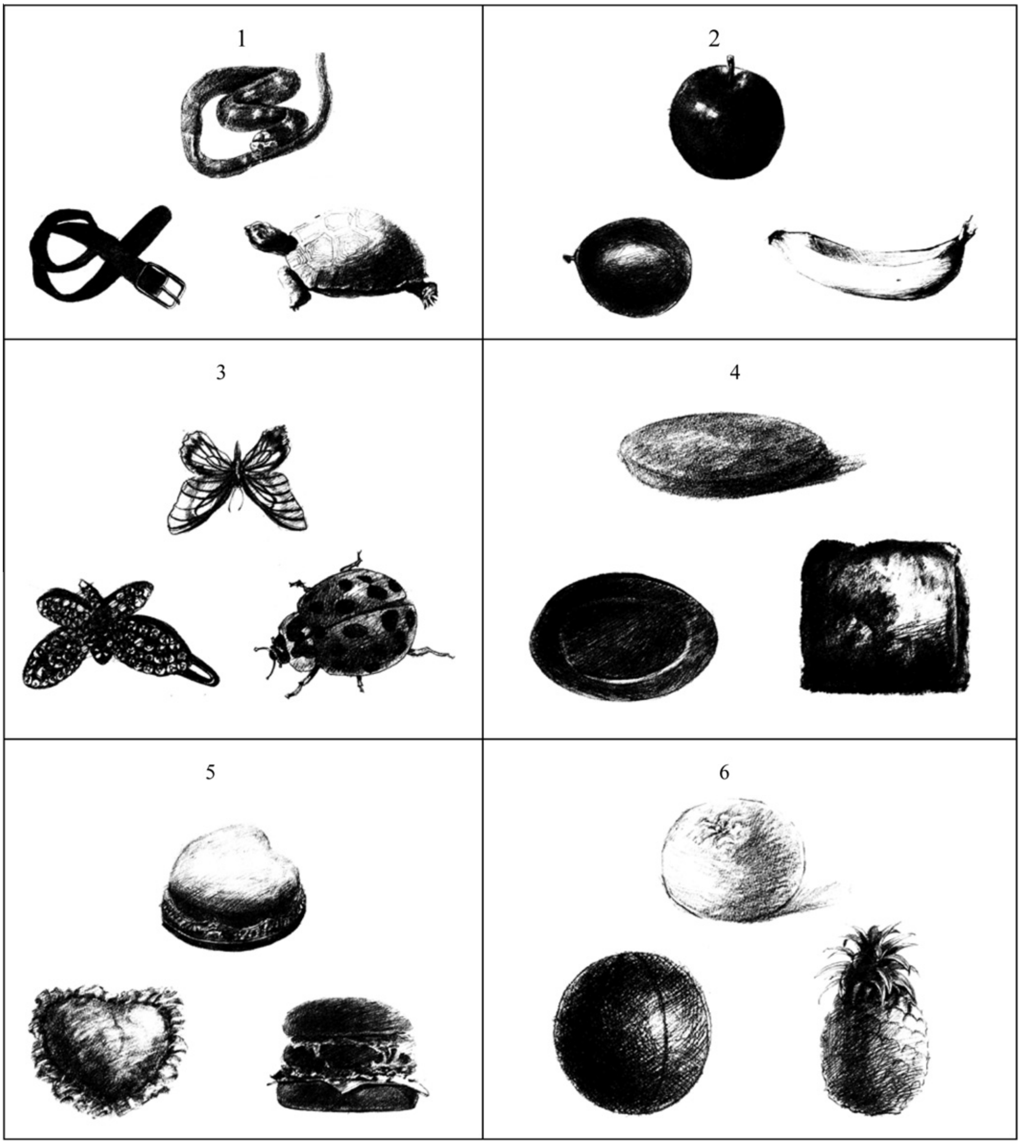

Fig. 3. Triads in Experiment 2.

Appendix B

See Fig. 3.

Appendix C

See triads in Fig. 4.

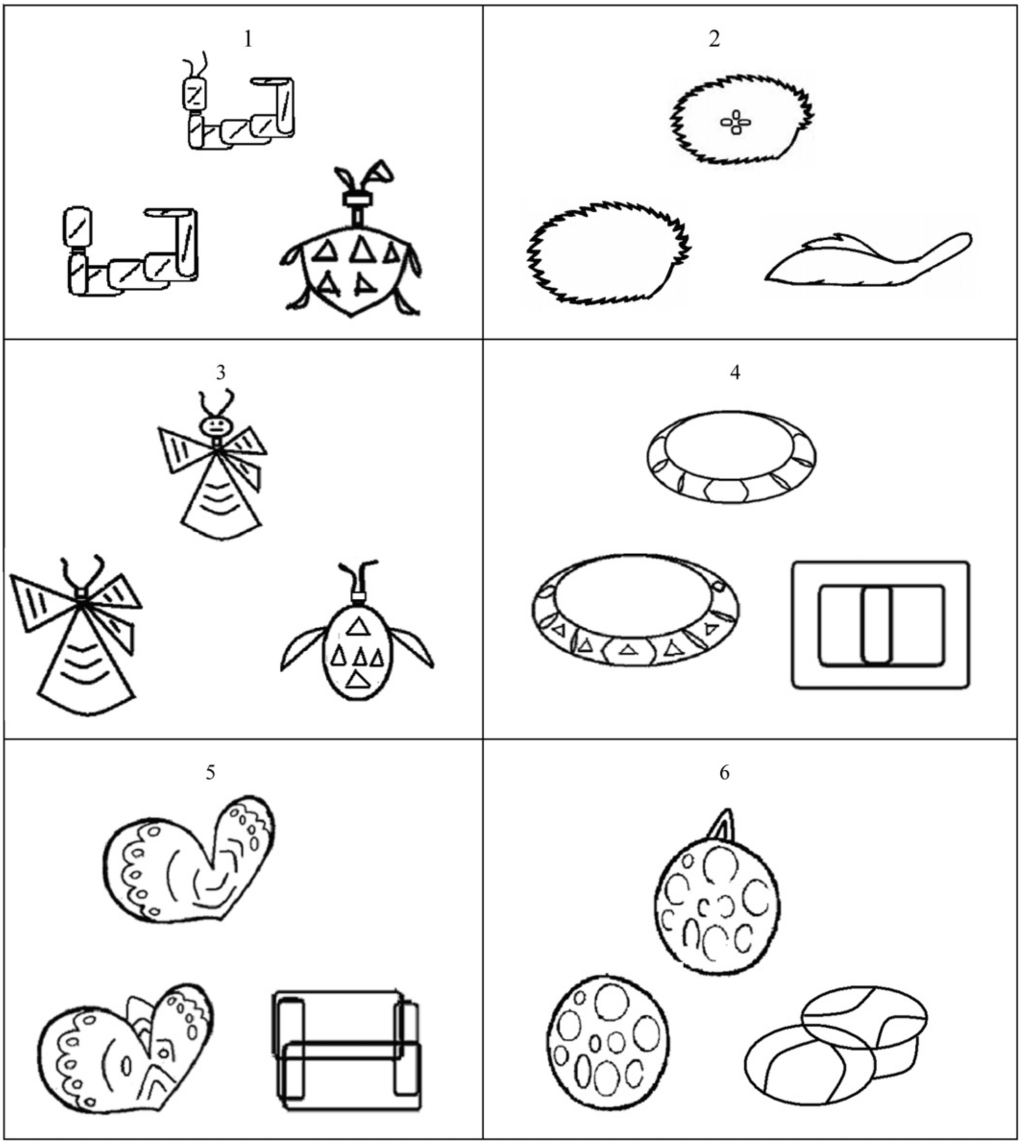

Fig. 4. Triads in Experiment 3.

References

Booth, A. E., Waxman, S. R., & Huang, Y. T. (2005). Conceptual information permeates word learning in infancy. *Developmental Psychology, 41*, 491–505.

Cimpian, A., & Markman, E. M. (2005). The absence of a shape bias in children's word learning. *Developmental Psychology, 41*, 1003–1019.

Daehler, M. W., Lonardo, R., & Bukatko, D. (1979). Matching and equivalence judgments in very young children. *Child Development, 50*, 170–179.

Davidson, N. S., & Gelman, S. A. (1990). Inductions from novel categories: The role of language and conceptual structure. *Cognitive Development, 5*, 157–176.

Deák, G., & Bauer, P. J. (1995). The effects of task comprehension on preschoolers' and adults' categorization choices. *Journal of Experimental Child Psychology, 60*, 393–427.

Deák, G., & Bauer, P. J. (1996). The dynamics of preschoolers' categorization choices. *Child Development, 67*, 740–767.

Farrar, M. J., Raney, G. E., & Boyer, M. G. (1992). Knowledge, concepts, and inferences in childhood. *Child Development, 63*, 673∼691.
Fisher, A. V. (2010). What's in the name? Or how rocks and stones are different from dogs and puppies. *Journal of Experimental Child Psychology, 105*, 198–212.
Fisher, A. V., Matlen, B., & Godwin, K. E. (2011). Semantic similarity of labels and inductive generalization: Taking a second look. *Cognition, 118*, 432–438.
Florian, J. E. (1994). Stripes do not a zebra make, or do they? Conceptual and perceptual information in inductive inference. *Development Psychology, 30*, 88–101.
Ganea, P. A., Pickard, M. B., & DeLoache, J. S. (2008). Transfer between picture books and the real world by very young children. *Journal of Cognition and Development, 9*, 46–66.
Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20*, 65–95.
Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8*, 404–409.
Gelman, S. A. (2009). Learning from others: Children's construction of concepts. *Annual Review of Psychology, 60*, 115–140.
Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a Dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology, 26*, 796–804.
Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183–209.
Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58*, 1532–1541.
Gelman, S. A., & O'Reilly, A. W. (1988). Children's inductive inferences within super-ordinate categories: The role of language and category structure. *Child Development, 59*, 876–887.
Gelman, S. A., & Waxman, S. R. (2007). Looking beyond looks: Comments on Sloutsky, Kloos, and Fisher, "When looks are everything: Appearance similarity versus kind information in early induction". *Psychological Science, 18*, 554–555.
Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development, 75*, 409–427.
Hayes, B. K., Heit, E., & Swendsen, H. (2010). Inductive reasoning. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 278–292.
Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development, 9*, 45–75.
Jaswal, V. K. (2004). Don't believe everything you hear: Preschoolers' sensitivity to speaker intent in category induction. *Child Development, 75*, 1871–1885.
Johnson, K. E., Scott, P., & Mervis, C. B. (1997). Development of children's understanding of basic–subordinate inclusion relations. *Developmental Psychology, 33*, 745–763.
Jones, S. S., & Smith, L. B. (1993). The place of perception in children's conception. *Cognitive Development, 8*, 113–139.
Landau, B., Smith, L. B., & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*, 299–321.
Mak, B. S. K., & Vera, A. H. (1999). The role of motion in children's categorization of objects. *Cognition, 71*, B11–B21.
Massey, C. M., & Gelman, R. (1988). Preschoolers' ability to decide whether a photographed unfamiliar object can move itself. *Developmental Psychology, 24*, 307–317.
Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89–115.
Myers, J. (2000). Rules vs. *Analogy in Mandarin classifier selection. Language and Linguistic, 12*, 187–209.
Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439.
Saalbach, H., & Imai, M. (2006). Categorization, label extension, and inductive reasoning in Chinese and German preschoolers: Influence of a classifier system and universal cognitive constraints. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th Conference of the Cognitive Science Society* (pp. 703–708). Mahwah, NJ: Lawrence Erlbaum.
Sloutsky, V. M. (2003). The role of similarity in the development of categorization. *Trends in Cognitive Sciences, 7*, 246–251.
Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General, 133*, 166–188.
Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science, 18*, 179–185.
Sloutsky, V. M., & Lo, Y.-F. (1999). How much does a shared name make things similar? 1: Linguistic labels and the development of similarity judgment. *Developmental Psychology, 35*, 1478–1492.
Sloutsky, V. M., Lo, Y.-F., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels and the development of inductive inference. *Child Development, 72*, 1695–1709.
Sloutsky, V. M., & Napolitano, A. C. (2003). Is a picture worth a thousand words? Preference for auditory modality in young children. *Child Development, 74*, 822–833.
Smith, C. L. (1979). Children's understanding of natural language hierarchies. *Journal of Experimental Child Psychology, 27*, 437–458.
Welder, A. N., & Graham, S. A. (2001). The influence of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development, 72*, 1653–1673.