Domain Differences in the Weights of Perceptual and Conceptual Information in Children's Categorization

Gil Diesendruck and Shimon Peretz
Online First Publication, February 25, 2013. doi: 10.1037/a0032049

CITATION
Domain Differences in the Weights of Perceptual and Conceptual Information in Children’s Categorization

Gil Diesendruck and Shimon Peretz
Bar-Ilan University

Visual appearance is one of the main cues children rely on when categorizing novel objects. In 3 studies, testing 128 3-year-olds and 192 5-year-olds, we investigated how various kinds of information may differentially lead children to overlook visual appearance in their categorization decisions across domains. Participants saw novel animals or artifacts of varying degrees of similarity to target categories and were asked to place them in 1 of 2 categories. Manipulated across studies was the kind of information pitted against visual similarity: internal information (Study 1, both 3- and 5-year-olds), intentional information (Study 2, both 3- and 5-year-olds), or labels (Study 3, only 5-year-olds). Overall, we found that for 5-year-olds, but not so for 3-year-olds, internal information had a stronger effect on the categorization of animals than of artifacts. Intentional information, in turn, had a stronger effect on both age groups’ categorization of artifacts than of animals. Labels too had a stronger effect on 5-year-olds’ categorization of artifacts than of animals. These findings are consistent with a domain-specific account of categorization, according to which the weight of different kinds of information on categorization decisions depends on children’s developing understanding of domains.

Keywords: categorization, domain specific, conceptual information, labels

Two broad views of object categorization have dominated the cognitive developmental literature for the past few decades. According to one perspective, categorization is driven by domain-general mechanisms. In particular, this perspective emphasizes the roles of similarity between objects and of statistical regularities in the environment, as guides to children’s decisions about how to group objects into categories (Sloutsky, Kloos, & Fisher, 2007; Smith, 2005). An alternative perspective argues that categorization is driven by domain-specific mechanisms (Hirschfeld & Gelman, 1994). The claim is that children hold naïve theories about different objects in the world (e.g., animals vs. artifacts), and these theories inform children’s categorization decisions (Carey, 1995; Gelman, 2003; Keil, 1989; Mandler, 2004).

One method adopted by researchers to adjudicate between these theoretical positions has been to pit “perceptual” information—that is, information available from simply seeing an object—against “conceptual” information—that is, information deriving from one’s naïve understanding of the target objects. For instance, in these experiments, children had to decide on the category membership of a target object that matched one object in terms of perceptual similarity, but matched another object in terms of conceptual similarity (e.g., having the same function; see, for instance, Gelman & Markman, 1986; Kemler Nelson, Russell, Duke, & Jones, 2000; Smith, Jones, & Landau, 1996). This line of work generated important insights about the kinds of information that children can take into account when making categorization decisions. However, it left the theoretical debate far from settled (see, for instance, the exchange between Gelman & Waxman, 2007; Sloutsky et al., 2007).

A different research strategy is to look at the possible interaction between perceptual and conceptual cues in determining children’s categorization decisions. And indeed, scholars from both sides of the theoretical divide have embraced such a strategy. For instance, from a domain-specific perspective, Davidson and Gelman (1990) found that whereas labels affect children’s categorization decisions about novel animals, perceptual similarity constrained this effect; that is, when the same label was applied to a visually diverse set of exemplars, children did not succeed in using the label to form a category. In a different vein, Booth and Waxman (2002) found that conceptual information (i.e., describing items as animals or artifacts) led children to consider different kinds of visual cues in the definition of categories. Complementarily, researchers from a domain-general perspective have also looked at the possible interaction between perceptual similarity and primarily labels in children’s categorization and induction. For instance, Sloutsky and colleagues have argued that a label is an added feature to any object, and as such, it contributes to the perceived overall similarity among identically labeled objects (Sloutsky & Fisher, 2004, 2011; Sloutsky, Lo, & Fisher, 2001).

Following this “interactional” perspective, the theoretical debate can be reframed in terms of the relative weight of the various sources of information to children’s categorization decisions. In particular, the questions become: (a) To what extent can conceptual information override perceptual information in determining children’s categorization decisions? (b) Does perceptual similarity
impose certain limits on categorization decisions? And most importantly, (c) do the kinds of effective conceptual information and their interactions with perceptual similarity vary across domains and with development? We address these questions in the present study.

The basic structure of all three studies was identical. Children were exposed to target items—either animals or artifacts—that varied in their degree of perceptual similarity to two candidate categories: from very low to very high similarity (see Figure 1). Crucially, target items always shared some conceptual information with the candidate categories. The questions were how much perceptual dissimilarity children were willing to overlook in order to categorize items on the basis of their conceptual similarity, whether different kinds of conceptual information affected this dynamics, and whether there were differences across domains and development in the manifestation of these interactions between conceptual and perceptual information.

The kinds of conceptual information assessed in the studies were ones previously identified as potentially having distinctive effects on children’s categorization of animals and artifacts. Specifically, in Study 1, we focused on the effect of information about internal properties on children’s categorization. As is discussed below, this is information that has been argued to affect children’s categorization of animals more so than that of artifacts. In contrast, in Study 2, we assessed the effect of information about creator’s intent on children’s categorization. This information arguably plays a more central role on children’s categorization of artifacts than of animals. Given that these two types of information are inherently linked to how children conceptualize the two domains—conceptualizations that may change across development—these two studies included children from two age groups: 3- and 5-year-olds. The last type of information evaluated was labels (Study 3). As is discussed below, whereas labels per se are not believed to intrinsically differentiate between the two domains, they may have differential effects given the supposed conceptual construal of the domains.

The general domain-specific hypothesis was that whereas information about internal properties (Study 1) would have a stronger effect on children’s categorization of animals than of artifacts, the reverse domain difference would obtain for information about creator’s intent (Study 2) and labels (Study 3). A domain-general position would not predict differences between domains in any study.

**Study 1: The Effect of Internal Information**

It has been argued that whereas animal categories are defined by inherent nonobvious properties causally responsible for animals’ physical appearances and behaviors (Atran, 1990; Gelman, 2003), artifact categories are defined by reference to what people intend an artifact to be, or by how people conventionally construe an artifact (Bloom, 1996; Keil, 1995). In other words, internal properties are at the core of people’s concepts of animal kinds.

Developmental studies reveal a somewhat complex picture regarding this differentiation between domains. On the one hand, Sobel and colleagues found a development increase from 3- to 4-year-olds in the extent to which they treat information about internal properties of artifacts as critical for determining category membership (Sobel & Corriveau, 2010; Sobel, Yoachim, Gopnik, Meltzoff, & Blumenthal, 2007). On the other hand, it seems that between 3 and 5 years of age, children start differentiating between the domains in terms of the conceptual centrality of internal properties. For instance, 3-year-olds are not as good as 4- and 5-year-olds in differentiating the insides of animals and machines (Gottfried & Gelman, 2005), and there are slight developmental changes from 3 to 4 years of age in children’s attribution of animals’, but not artifacts’, movements as caused by internal properties (Gelman & Gottfried, 1996). Moreover, by 4 years of age—but less so at 3 years of age—children recognize the importance of insides for animal category membership (Gelman & Wellman, 1991). Indeed, by 4 years of age, children generalize novel internal properties based on animal category membership (Gelman & Markman, 1986), and do so to a lesser extent based on artifact category membership (Gelman, 1988; Gelman & O’Reilly, 1988). Complementarily, 4-year-olds regard internal properties as more definitional of animal than of artifact categories (Diesendruck, Gelman, & Lebowitz, 1998), and they recognize that removing the insides of an object affects its identity more than removing its surface properties (Gelman & Wellman, 1991). By 8 years of age, children believe that external transformations change artifacts’, but not animals’, category membership (Keil, 1989).

Taken together, these findings suggest that whereas internal properties may be central to children’s conceptualization of both domains, between 3 and 5 years of age, its weight in the conceptualization of animals becomes greater than in the conceptualization of artifacts. In Study 1, we investigated how 3- and 5-year-olds treat internal properties when assigning animal or artifact categories, and whether internal property information allowed children to overcome perceptual dissimilarity in these assignments, with equal force across domains.

**Method**

**Participants.** Sixty-four 3-year-olds ($M_{age} = 38.14$ months, $SD = 5.42$; 27 boys and 37 girls) and 64 5-year-olds ($M_{age} = 62.55$ months, $SD = 5.19$; 36 boys and 28 girls) participated in this study. An additional group of 16 5-year-olds participated in the calibration study, used to select the experimental stimuli. Children were monolingual Hebrew speakers, recruited from several Israeli kindergartens. Only children with signed parental permission participated. All children received a sticker for their participation.

**Design.** The experiment had a mixed design, with domain (animals vs. artifacts) and condition (internal information vs. no information) as between-subjects factors and similarity ratios (.12, .37, .62, .87) as a within-subjects variable. Participants were assigned randomly to one of four groups, resulting from the crossing of the two between-subjects factors: domain and condition.

**Materials.** Stimuli consisted of 11 cm × 11 cm pictures of animals and artifacts, presented on a 13-in. laptop screen. The animal sets consisted of exotic animals likely to be unfamiliar to young children. The artifact sets consisted of exotic animals likely to be unfamiliar to young children. The artifact sets consisted of parts of, or modified, atypical objects. There were four sets of animal category pairs and four sets of artifact category pairs (see Table 1 for the full list). Each set consisted of two contrasting target categories (e.g., cranes vs. chameleons) and four test stimuli. The target categories depicted a group of three identical animals or artifacts from a given category. Test stimuli consisted of a picture of a single animal or artifact, which had a different degree of similarity to the set’s
Figure 1. Examples of animal (left column) and artifact (right column) sets at the different similarity ratios, used in all three studies. The similarity ratios are between the test item and Target Category A. Depicted here is only one of four triads from each set.
target categories. In order to minimize intervention of former knowledge, we pretested children to make sure that all animals and artifacts were unfamiliar to most of the participants. Only one child identified one of the animal categories (the lemurs).

The sets were created as follows. First, a pair of animal or artifact pictures from different categories was used to create a 16-step sequence, in which one member of a pair was morphed into another member (e.g., a crane into a chameleon). These morphing sequences were created using Abrosoft FantaMorph 3.7.1 software. Out of the 16 pictures in a sequence, we selected every second picture to serve as a “candidate” test stimulus in the actual experiments (total of eight pictures per set). We repeated this procedure for four pairs of animals and four pairs of artifact categories, creating 64 triads composed of two target category exemplars and one candidate test stimulus (2 domains × 4 pairs of categories × 8 candidate test stimuli).

These triads were subjected to a calibration study to quantify the similarity of each candidate test stimulus to each one of the target exemplars. This calibration study followed the method used by Sloutsky and Fisher (2004). A group of 16 5-year-olds were presented with the 64 triads, one at a time, and were asked whether the candidate test stimulus looked more like the exemplar from Target Category A or more like the exemplar from Target Category B. For each triad, the number of choices of Target Category A divided by the number of participants was the measure of similarity between the candidate test stimulus and Target Category A. For instance, if only two out of the 16 5-year-olds of a given test stimulus was more similar to the exemplar from Target Category A than to that from Target Category B, then that stimulus received a similarity ratio of .12 to Category A. On the basis of this calibration study, 32 triads were selected, representing four “similarity” ratios of test stimuli to Target Category A: .12, .37, .62, and .87 (see Figure 1 for examples of various stimuli, and Figure 2 for a full set). At the end of the calibration study, each animal and artifact set had one test stimulus at each of the four ratios of similarity.

Procedure. Each child was interviewed individually in a quiet room in his or her kindergarten. The entire procedure was conducted using a laptop computer for presentation of the pictures. In all groups, participants first saw a picture of a detective child named Danny. They were told:

Look, this is Danny. Danny is a detective kid. He loves helping [lost animals/putting objects] back to where they belong. Sometimes Danny doesn’t know where to return the [animals/objects]. Would you like to help Danny return the [animals/objects] to where they belong?

The first triad was then presented. The experimenter revealed the two target categories to the participant by saying, for instance, “Look at these [animals/objects] [pointing to Target A, which appeared on one side of the screen] they have a small [brain/motor] inside of them. I’ll show you one of them up close.” The experimenter clicked on Target A, making two of the exemplars in Target A move to the background and one of them move to the foreground, resizing to be the same size as the test stimuli to appear later on. The experimenter then continued: “Now look at these [animals/objects] [pointing to Target B], they have a large [brain/motor] inside of them. I’ll show you one of them up close,” and proceeded just as was done with Target A.

The experimenter then revealed one of the test stimuli for the set, shown at the bottom center of the screen, equidistant from the two target categories. In the internal information condition, the experimenter pointed to the test stimulus and said: “Look at this one, it has a small [brain/motor]. Can you help Danny return it to the group of the same kind?” In all trials, Target Category A had the same internal property as the test stimulus. Consequently, low-similarity test stimuli (.37, and especially .12) had the same internal properties as the least similar target category, thus creating a conflict between internal properties and degree of similarity. In trials of high similarity (.62, and especially .87), internal properties and similarity cohered, as the properties were applied to the category most similar to the test stimulus.

In the no-information condition, the experimenter pointed to the test stimulus and said: “Look at this one. Can you help Danny return it to the group of the same kind?” Evidently, in this condition, there was no conflict between internal information and physical similarity.

After recording the participant’s response (either verbal or pointing), the next test stimulus pertinent to the set was shown, and the experimenter repeated his question. After all four test stimuli for a given set were shown, the next set was presented. Different internal organs and parts were used for each set, but the pairings of properties to sets were the same for all participants (see Table 1). Altogether, every participant completed 16 trials: either four trials for four different animal sets or four trials for four different artifact sets. Both order of sets and of test stimuli were randomized for each participant, and so was the introduction and placement of the categories. No feedback was given to the participants on their performance.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Set</th>
<th>Internal Property A</th>
<th>Internal Property B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animals</td>
<td>Northern tamandua–armadillo</td>
<td>Green blood</td>
<td>Yellow blood</td>
</tr>
<tr>
<td></td>
<td>Booby seabird–sifaka lemur</td>
<td>Two-parts heart</td>
<td>Three-parts heart</td>
</tr>
<tr>
<td></td>
<td>Sandhill crane–calumma chameleon</td>
<td>Small brain</td>
<td>Large brain</td>
</tr>
<tr>
<td></td>
<td>Southern tamandua–pangolin</td>
<td>5 bones</td>
<td>10 bones</td>
</tr>
<tr>
<td></td>
<td>Part of candle–part of fishing tool</td>
<td>Two-parts screw</td>
<td>Three-parts screw</td>
</tr>
<tr>
<td></td>
<td>Handle of electric tool–rolled-up elastic band</td>
<td>5 rods</td>
<td>10 rods</td>
</tr>
<tr>
<td></td>
<td>Part of Japanese lamp–casino chips stuck together</td>
<td>Green oil</td>
<td>Yellow oil</td>
</tr>
<tr>
<td></td>
<td>Part of olive oil tool–leveling tool attached to wooden frame</td>
<td>Small motor</td>
<td>Large motor</td>
</tr>
</tbody>
</table>

Note. Items are listed in the order that they appear in Figure 1, from top to bottom.
Results

Participants’ responses were scored according to the number of trials at each ratio of similarity in which they selected Target A. These scores, which ranged from 0 (i.e., selected Target A on none of the trials) to 4 (i.e., selected Target A on all trials) were used as dependent measures in all analyses. A repeated measures analysis of variance (ANOVA), which included age group (3s vs. 5s), domain (animals vs. artifacts), and condition (internal vs. no information) as between-subjects factors, and similarity ratio (.12, .37, .62, and .87) as a within-subjects factor, revealed the following significant effects: similarity ratio, $F(3, 118) = 125.04, p < .001, \eta^2 = .76$; domain, $F(1, 120) = 14.01, p < .001, \eta^2 = .11$; and condition, $F(1, 120) = 126.10, p < .001, \eta^2 = .51$; and two-way interactions between domain and similarity ratio, $F(3, 118) = 3.70, p < .05, \eta^2 = .09$, and condition and similarity ratio, $F(3, 118) = 14.19, p < .001, \eta^2 = .27$. Finally, these effects were subsumed under two theoretically central three-way interactions: among domain, condition, and similarity ratio, $F(3, 118) = 4.59, p < .005, \eta^2 = .10$, and among age group, condition, and similarity ratio, $F(3, 118) = 3.57, p < .05, \eta^2 = .08$. There were no other significant effects. We followed up on these interactions by analyzing each age group separately.

Analyses on 3-year-olds. A repeated measures ANOVA similar to the one described above revealed a significant main effect of similarity ratio, $F(3, 58) = 37.66, p < .001, \eta^2 = .66$, with children being more likely to place test items with the most similar target category. There was also a significant interaction between similarity ratio and condition, $F(3, 58) = 4.69, p < .01, \eta^2 = .19$. A follow-up multivariate analysis of variance (MANOVA) revealed that at all levels of similarity, 3-year-olds were more likely to assign the test item to Target Category A in the internal information condition than in the no-information condition (all $ps < .005$). The interaction likely resulted from the degree of difference between conditions at each ratio of similarity. Most importantly, the ANOVA revealed no other significant effects, including the theoretically central interactions involving both domain and condition. In sum, 3-year-olds’ categorization decisions were swayed by both perceptual similarity and information about internal parts, irrespective of the domain of the entities being categorized.

Analyses on 5-year-olds. The repeated measures ANOVA revealed the following significant effects: similarity ratio, $F(3, 58) = 115.82, p < .001, \eta^2 = .86$; condition, $F(1, 60) = 82.60, p < .001, \eta^2 = .58$; and domain, $F(1, 60) = 14.80, p < .001, \eta^2 = .20$; and two-way interactions between domain and similarity ratio, $F(3, 58) = 2.78, p < .05, \eta^2 = .13$, and between condition and similarity ratio, $F(3, 58) = 16.88, p < .001, \eta^2 = .47$. Differently from the findings with 3-year-olds, these effects were all subsumed under the theoretically central three-way interaction among domain, condition, and similarity ratio, $F(3, 58) = 3.31, p < .05, \eta^2 = .15$. Figure 3 displays the relevant data. There were no other significant effects.

We followed up on the above interaction by first conducting separate MANOVAs looking at the effect of condition in each domain. In the domain of animals, the omnibus MANOVA revealed a significant effect of condition, $F(4, 27) = 17.02, p < .001, \eta^2 = .72$, which held true at all levels of similarity except at .87 ($ps < .01$). In general, children were more likely to assign the test animal to Target Category A in the internal than in the no-information condition. In the domain of artifacts, the omnibus MANOVA also revealed a significant effect of condition, $F(4, 27) = 8.32, p < .001, \eta^2 = .55$, which held true at all levels...
Discussion

Study 1 showed that, in general, internal property information affected both age groups’ categorization decisions in both domains. However, whereas 5-year-olds were more influenced by this type of information when categorizing animals than when categorizing artifacts, the effect on 3-year-olds was equivalent in the two domains. This developmental difference seems in line with the studies reviewed earlier, showing that between the ages of 3 and 4, children become more sensitive to the role of insides in the definition of animals, as opposed to artifacts, category membership (e.g., Diesendruck et al., 1998; Gelman & Wellman, 1991).

The finding that internal information affected the categorization of novel artifacts—in both 3- and 5-year-olds—is somewhat surprising, as it seems to contradict previous results (e.g., Diesendruck et al., 1998). One possible conciliation of these findings is that whereas Diesendruck et al. asked children to categorize familiar artifacts (e.g., cars), here children were asked to categorize unfamiliar artifacts. It is possible that when categorizing familiar artifacts, children rely less on internal properties than on well-known functional and intentional properties. In turn, upon encountering novel artifacts, children might be more tolerant of the kind of information they consider when deciding on the kind of an artifact (see, for instance, Gelman & Wellman, 1991; Sobel & Corriveau, 2010; Sobel et al., 2007).

As for the domain differences in the effectiveness of internal property information among 5-year-olds, the finding is consistent with previous studies on this matter (Diesendruck et al., 1998; Gelman, 1988; Gelman & O’Reilly, 1988). The present design, in which internal property information was pitted against varying degrees of perceptual similarity, allowed us to discern in more detail the relative weights of these sources of information. In particular, the findings revealed that the effect of internal property information on 5-year-olds’ assignment of animal categories was quite dramatic. For instance, for animals rated as 12% similar to Category A, children in the no-information condition assigned the test stimuli to Category A approximately 3% of the time; that figure was approximately 35% for animals in the 37% similarity ratio. In turn, the corresponding figures for children in the internal information condition were over 75% and 90%, respectively. (In comparison, whereas the figures for artifacts in the no-information condition were similar to those reported above for animals, the corresponding figures in the internal information condition were 40% and 65%.) In other words, internal property information about animals led children to radically switch their default categorization decisions, and almost completely ignore the degree of perceptual dissimilarity between the test stimuli and the target category.

Taken together, these findings attest that by 5 years of age, internal property information has a more significant weight in children’s categorization of animals than of artifacts—a conclusion consistent with a domain-specific account of categorization. In Study 2, we tested whether the opposite domain difference is true when providing children with information about the intentions of the creator of animals and artifacts.

Study 2: The Effect of Intentional Information

Studies indicate that children’s definitions of artifact categories are primarily influenced by functional or intentional properties. In particular, Bloom (1996) proposed that children decide what an artifact is, based on intuitions about what it was intended to be by its creator. In fact, and differently from what has been found regarding internal properties, it seems that children reach this
understanding fairly early in development. Specifically, by 3 years of age, children take intentional information, or information about the function of artifacts, as critical for defining the category an artifact belongs to (Asher & Kemler Nelson, 2008; Diesendruck, Markson, & Bloom, 2003; Gelman & Bloom, 2000; Jaswal, 2006; cf. Smith et al., 1996). Notably, this understanding undergoes sophistication with development. For instance, 4-year-olds are more likely than 3-year-olds to rely on an object’s creator as a source for the object’s name (Sabbagh & Baldwin, 2001) and are better at understanding the correspondence between intended design and artifact structure (Kelemen, Seston, & Saint Georges, 2012). Also, 6-year-olds are more likely than 4-year-olds to rely on an object’s original intended function as a determinant of category membership (Matan & Carey, 2001).

To our knowledge, no study has directly assessed the effect of “creator’s intent” on children’s definition of animal categories. Thus, it is unclear whether and when children differentiate between the two domains in this regard. In Study 2, we investigated this matter by adopting the design of Study 1, but instead of looking at the effect of internal property information, we assessed the effect of information about a creator’s intent in children’s assignment of artifact and animal categories. The hypothesis regarding domain differences was exactly the opposite of that posited in Study 1. Namely, intentional information should be more effective in determining children’s categorization of artifacts than of animals. Given the evidence for children’s early sensitivity to intentional information when categorizing artifacts, we hypothesized that domain differences might appear even among 3-year-olds.

Method

Participants. Sixty-four 3-year-olds (Mage = 40.25 months, SD = 4.81; 33 boys and 31 girls) and 64 5-year-olds (Mage = 64.8 months, SD = 4.2; 35 boys and 29 girls) participated in this study. Recruitment procedure and demographic information were identical to those of Study 1. None of the children had participated in Study 1.

Design. The experiment had the same mixed design as Study 1, with domain (animals vs. artifacts) and condition (intentional information vs. no information) as between-subjects factors and similarity ratio (.12, .37, .62, and .87) as a within-subjects variable. Participants were assigned randomly to one of four groups, resulting from the crossing of the two between-subjects factors: domain and condition.

Materials. Materials were the same as those used in Study 1.

Procedure. The procedure started off exactly as in Study 1. Only here, the boy character in the story—“Danny”—was introduced as being “a scientist, who likes to invent new [animals/objects].” The experimenter introduced the target categories to the participant using an artificial two-syllable label: “Look, here are Musos [pointing to Target Category A, which appeared on one side of the screen]. I’ll show you one of them up close.” After showing Target Category A by enlarging one of its exemplars, the experimenter introduced Target Category B in the same manner, for example, “And here are Gagas [pointing to Target Category B, which appeared on the opposite side of the screen]. I’ll show you one of them up close.”

After introducing the target categories, the story continued in different ways according to the child’s condition. In the intentional condition, Danny said: “I want to create a Muso, just like these ones [pointing to Target Category A].” In the no-information condition, Danny said: “I want to create something.” The story described how Danny went to his laboratory to work on his plan. Then the experimenter presented the test stimulus and said to the child, “Hey, look what Danny made. Can you put it with the group of the same kind? Should it go with the Gagas or with the Musos?” In all trials of the intentional condition, the test stimuli were intended to be like the items in Target Category A. Thus, in trials of low-similarity ratios, intentional information and similarity provided conflicting cues about the category membership of the test stimulus. In the no-information condition, there were no conflicting cues, as children only had similarity as a cue to category membership.

After writing down the participant’s response, the next test stimulus in the set was shown. After all four test stimuli of a set were shown, the next set was presented. The names used for Target Categories A and B, respectively, were Lulu and Tipi, Gaga and Muso, Pufi and Ripa, and Lupa and Zizi. Altogether, every participant completed 16 trials: either four trials for four different animal sets or four trials for four different artifact sets. Presentation orders were counterbalanced as in Study 1.

Results

Participants’ responses were scored as in Study 1. These scores were used as the dependent measure in a repeated measures ANOVA, which included age group (3s vs. 5s), domain (animals vs. artifacts), and condition (intentional information vs. no information) as between-subjects factors and similarity ratio (.12, .37, .62, and .87) as a within-subjects factor. The ANOVA revealed the following significant effects: similarity ratio, F(3, 118) = 273.35, p < .001, η² = .87, and condition, F(1, 120) = 10.43, p < .005, η² = .08; and two-way interactions between condition and domain, F(1, 120) = 9.41, p < .005, η² = .07; domain and similarity ratio, F(3, 118) = 3.47, p < .05, η² = .08; condition and similarity ratio, F(3, 118) = 8.66, p < .001, η² = .18; and age group and similarity ratio, F(3, 118) = 34.60, p < .001, η² = .47. Finally, these effects were subsumed under the three-way interaction among domain, condition, and similarity ratio, F(3, 118) = 7.01, p < .001, η² = .15. Figure 4 displays this interaction. There were no other significant effects. Given that age group was not a factor in the three-way interaction, it was excluded from the follow-up analyses.

As in Study 1, we followed up on the above interactions by first conducting separate MANOVAs looking at the effect of condition in each domain separately.

In the domain of animals, the omnibus MANOVA revealed that the effect of condition was not significant, F(4, 59) = 0.74, p > .5, η² = .05, and this was true at each ratio of similarity. In other words, providing children with information about the intentions of the creator of an animal did not affect children’s assignment of category membership. In turn, in the domain of artifacts, the omnibus MANOVA revealed a significant effect of condition, F(4, 59) = 9.78, p < .001, η² = .40, which held true primarily at the .37 (p < .001) and .62 (p = .076) ratios of similarity. As can be seen in Figure 4, in general children were more likely to assign
the test artifact to Target Category A in the intentional than in the no-information condition.

For the sake of equivalence across studies, we also conducted MANOVAs looking at the effect of domain in each condition separately—as was done in Study 1. In the no-information condition, the omnibus MANOVA revealed a significant effect of domain, $F(4, 59) = 6.08, p < .001, \eta^2_p = .29$, which held true primarily at the .37 ($p < .005$) and .62 ($p < .05$) ratios of similarity. This unexpected finding indicates that children were more likely to assign animals, compared with artifacts, to Target Category A. In turn, in the intentional information condition, the omnibus MANOVA revealed no significant effect of domain, $F(4, 59) = 1.87, p = .13, \eta^2_p = .11$. Nonetheless, the univariate ANOVAs revealed a significant effect of domain on children’s assignment of category membership at the .37 similarity ratio ($p < .05$). Crucially, as can be seen in Figure 4, the direction of the effect here was the opposite of the one found in the no-information condition. Namely, children in the intentional condition were more likely to categorize artifacts, compared with animals, as belonging to Target Category A. In other words, intentional information led children to overcome their perceptually based bias regarding domain differences.

**Discussion**

Study 2 revealed that whereas intentional information had no effect on children’s categorization decisions of animals, it affected children’s decisions regarding artifacts’ category membership. Furthermore, the finding that age group did not interact in significant ways with domain and/or condition indicates that the above pattern held similarly for both 3- and 5-year-olds.

The finding that intentional information had no effect on 3- and 5-year-olds’ categorization of animals is consistent with the idea that children construe animal categories as natural kinds, that is, as objective groupings of natural reality impervious to cultural conventions (Kalish, 1998; Rhodes & Gelman, 2009). According to this interpretation, intentional information was discounted because it is viewed a priori as irrelevant for animal category assignment. Alternatively, it is important to point out that children may have a tendency to construe nature—animals categories included—in teleological terms (Kelemen, 2003), and thus may be willing to accept an intentional agent as the creator of such categories (Evans, 2000). Following this argument, it is possible that children in the present study were not swayed by the intentional information because they were unconvinced by the power of the specific intentional agent used in the present study—a boy—to define animal categories. In either case, children’s pattern of responses is indicative of domain-specific reasoning.

The finding that intentional information affected young children’s categorization of artifacts is consistent with previous studies (Asher & Kemler Nelson, 2008; Diesendruck et al., 2003; Gelman & Bloom, 2000) and supports Bloom’s (1996) notion that children categorize artifacts by tracking the intentions of artifact creators. More broadly, the finding complements those of Study 1 by showing that differently from internal information, intentional information plays a more significant role in the categorization of artifacts than of animals. In fact, the finding that children were more likely to categorize artifacts than animals according to intentional information indicates that the domain difference found in Study 1 among 5-year-olds did not derive from some general difference between domains in their susceptibility to conceptual information. Rather, each domain benefited from information about specific properties. This dynamic interaction between property type and domain is a direct prediction of a domain-specific account of categorization.

**Study 3: The Effect of Labels**

Although there are numerous developmental studies assessing the relation between labeling and categorization, most have not systematically addressed the possibility of domain differences (e.g., Waxman & Braun, 2005; Xu, 2002). Moreover, although domain-general similarity-based accounts have provided extensive evidence on the effect of labels on children’s categorization, they downplay the plausibility of finding a priori domain differences (Sloutsky & Fisher, 2011).

This relative “neglect” of investigations regarding domain differences vis-à-vis labels may be partly due to the notion that labels per se are not taken as constituent components of people’s naïve theories of the world. This stands in contrast, for instance, with the notions exposed before that internal properties are thought to be central to people’s naïve theories about animals, and intentional information is central to theories about artifacts. From a domain-
specific perspective, however, one might expect labels to have a differential effect on categorization across domains because of people’s beliefs about the very nature of the categories in the different domains. In particular, Diesendruck (2003) hypothesized that given that animal categories seem to be taken to capture natural discontinuities in the environment, and artifact categories are thought to be conventionally and arbitrarily created, the latter may be more susceptible to how categories are labeled than the former.

This differentiation is supported by a number of studies on both adults (Cimpian, Gelman, & Brandone, 2010; Diesendruck & Gelman, 1999; Estes, 2003) and young children (Kalish, 1998; Rhodes & Gelman, 2009). For instance, Rhodes and Gelman found that by 5 years of age, children believe that animal, but not artifact, categories are objectively true, and should thus be universally and normatively endorsed. In other words, whereas for animals, “physical reality”—as signaled by visual appearance—might constrain which categories are deemed acceptable (see, for instance, Davidson & Gelman, 1990), for artifact categories, physical reality is a lesser factor, and what matters most is what people conventionally decide (see, for instance, Diesendruck & Markson, 2011; Siegel & Callanan, 2007).

One study that provided some indirect evidence consistent with the above hypothesis was conducted by Jaswal and Markman (2007). In that study, 24-month-olds saw pictures of hybrid animals and artifacts, such that a target object looked like a member of one familiar category but was labeled as a member of a different familiar category (e.g., a hybrid of a cat and a dog that looked more like a cat, but was labeled “a dog”). Children were asked whether the target object would share the same property as the more similar category, or as the same-labeled category. Even though Jaswal and Markman found that, overall, children inferred based on the shared label, analyses against chance revealed that whereas children systematically drew inferences about artifacts based on the label, they did not do so for animals.

The present Study 3 followed a similar method to that used by Jaswal and Markman (2007) and Sloutsky and Fisher (2004), but with some important changes. First, in contrast to Jaswal and Markman, we used a categorization rather than an induction task. In other words, in the present studies, children were asked to determine to which category an item belongs, rather than which property a labeled item has. Second, in contrast to Sloutsky and Fisher, given our interest in how children form categories, we used as stimuli novel categories rather than familiar ones. This is an important difference because it required children to rely exclusively on the information available to them at the task rather than on their prior knowledge about the categories. Third and crucially, in Study 3, we systematically compared the effectiveness of labeling in driving children’s categorization in the two domains. Given the somewhat exploratory nature of the study, and the lack of clear hypotheses about developmental changes, Study 3 was conducted only with 5-year-olds.

Method

Participants. Participants were 64 5-year-olds ($M_{age} = 64.05$ months, $SD = 5.5$; 34 boys and 30 girls). Recruitment procedure and demographic information were identical to those of Study 1. None of the children had participated in Studies 1 or 2.

Design. The experiment had the same mixed design of the previous studies, with domain (animals vs. artifacts) and condition (label vs. no label) as between-subjects factors and similarity ratio (.12, .37, .62, and .87) as a within-subjects variable. Participants were assigned randomly to one of four groups, resulting from the crossing of the two between-subjects factors: domain and condition.

Materials. Materials were the same as those used in the previous studies.

Procedure. The procedure started off exactly as in Study 1, that is, introducing children to “Danny, the detective boy,” who sometimes needs help returning objects or animals to where they belong. The only difference from Study 1 had to do with how the experimenter introduced the target categories and test stimuli. Here, the experimenter introduced the target categories to the participant using an artificial two-syllable label: “Look, here are Musos [pointing to Target Category A, which appeared on one side of the screen]. I’ll show you one of them up close”; and then, “And here are Gagas [pointing to Target Category B, which appeared on the opposite side of the screen]. I’ll show you one of them up close.”

After introducing the target categories, a test stimulus appeared. In the label condition, the experimenter pointed to the test stimulus and labeled it with the same label previously applied to Target Category A. For instance: “This Muso is lost. Can you help Danny return it to the group of the same kind? Should it go with the Gagas or with the Musos?” In all trials, the test stimulus was given the same label as Target Category A. This means that in trials of low-similarity ratios, label and similarity provided conflicting cues about the category membership of the test stimulus, as the label referred to the category least similar to the test stimulus. In the no-label condition, the experimenter pointed to the test stimulus and simply said: “This one is lost. Can you help Danny return it to the group of the same kind? Should it go with the Gagas or with the Musos?” Thus, in the no-label condition, there were no conflicting cues, as children only had similarity as a cue to category membership.

After writing down the participant’s response, the next test stimulus pertinent to the set was shown. After all four test stimuli for a given set were shown, the next set was presented. The names used for Target Categories A and B were the same as the ones used in Study 2. Altogether, every participant completed 16 trials: either four trials for four different animal sets or four trials for four different artifact sets. Presentation orders were counterbalanced as in the previous studies.

Results

Participants’ responses were scored as in Studies 1 and 2. These scores were used as the dependent measure in a repeated measures ANOVA, which included domain (animals vs. artifacts) and condition (label vs. no label) as between-subjects factors and similarity ratio (.12, .37, .62, and .87) as a within-subjects factor. Figure 5 displays the relevant data. The ANOVA revealed the following significant effects: similarity ratio, $F(3, 58) = 800.00, p < .001$. $\eta^2 = .98$, and Condition, $F(1, 60) = 26.07, p < .001$, $\eta^2 = .30$; and two-way interactions between domain and condition, $F(1, 60) = 13.85, p < .001$, $\eta^2 = .19$, and between condition and similarity ratio, $F(3, 58) = 11.28, p < .001$, $\eta^2 = .37$. These
effects were subsumed under the theoretically central three-way interaction among domain, condition, and similarity ratio, $F(3, 58) = 7.03, p < .001$, $\eta^2 = .27$. There were no other significant effects.

As in the previous studies, we followed up on the interactions by conducting a MANOVA, looking at the effect of condition within each domain separately.

In the domain of animals, the omnibus MANOVA revealed that the effect of condition was not significant, $F(4, 27) = 1.74, p > .17$, $\eta^2 = .21$. Nonetheless, the univariate ANOVAs revealed that labels did affect children’s assignment of animal category membership in the .12 similarity ratio ($p < .05$). In other words, providing children with the labels of animals had only a minor effect on children’s assignment of category membership. In turn, in the domain of artifacts, the omnibus MANOVA revealed a significant effect of condition, $F(3, 28) = 17.49, p < .001$, $\eta^2 = .65$, which held true primarily at the .37 ($p < .001$) and .62 ($p < .05$) ratios of similarity. As can be seen in Figure 5, children were more likely to align the test artifact to Target Category A in the label than in the no-label condition.

The complementary MANOVAs looking at the effect of domain in each condition separately revealed that in the no-label condition, the effect of domain was not significant, $F(4, 27) = 2.23, p > .1$, $\eta^2 = .19$, which held true at all ratios of similarity. In turn, in the label condition, the effect of domain was significant, $F(4, 27) = 4.61, p < .01$, $\eta^2 = .41$, and held true primarily at the .37 ($p < .005$) and .62 ($p = .07$) ratios of similarity. As can be seen in Figure 5, at these intermediate levels of similarity, 5-year-olds were more likely to categorize labeled artifacts, compared with animals, as belonging to Target Category A.

Discussion

 Whereas previous studies provided comprehensive empirical bases to draw predictions about the direction of effects in Studies 1 and 2, this was not the case for Study 3. This was so because the relative weight of labels on children’s categorization across domains had not been systematically examined. The present findings reveal that for 5-year-olds, labels were weighted more heavily when determining the category membership of a novel artifact than of a novel animal.

As noted in the Results, labels did have a minor effect on children’s categorization decisions regarding animals, primarily when test stimuli were highly dissimilar to the same-labeled target category. This finding is consistent with previous results from both a domain-specific (e.g., Gelman & Markman, 1986) and domain-general (e.g., Sloutsky & Fisher, 2004) perspective. Nonetheless, labels had an even stronger effect on children’s categorization decisions regarding artifacts. To illustrate, for artifacts rated as 37% similar to Category A, children in the no-label condition assigned the test stimuli to Category A approximately 23% of the time. When labels were provided for these very same stimuli, this figure rose to over 83% of the time.

This domain difference has been hinted at in a study by Jaswal and Markman (2007) and is consistent with Diesendruck’s (2003) domain-specific hypothesis. It seems that children are less tolerant of labeling practices that violate how they perceive animal categories than they are of such practices when applied to artifacts.

General Discussion

The present three studies were designed to assess the relative weights of perceptual information and various types of conceptual information in children’s decisions about the category membership of animals and artifacts. Although there have been numerous studies on this topic, the present series is the first to address this issue by exhaustively assessing three different types of information (internal, intentional, and label), within the same task, in regard to the same stimuli, and by providing a gradient measure of perceptual similarity. The findings revealed novel and consistent patterns of differences between domains in the weights of the various types of information, especially among 5-year-olds.

First and foremost, it is clear that, especially in the absence of any other type of information, children’s categorization decisions quite systematically varied according to the degree of similarity in the appearance of test stimuli and target categories. Participants across all studies and domains were sensitive to the morphing manipulation, and succeeded in capturing even slight differences between test stimuli and target categories when deciding on category assignment. This vouches for the importance of this domain-general cue in children’s decisions about how to group any kind of entities into categories (Sloutsky & Fisher, 2004; Smith, 2005).
Second, however, it was also clear that 3-year-olds’, but especially 5-year-olds’, reliance on perceptual similarity as a cue to category assignment varied as a function of the domain of the entities being categorized and the type of conceptual information competing with perceptual similarity. In particular, Studies 1 and 2 showed that whereas information about the internal properties of the test stimuli had a stronger impact on 5-, but not 3-, year-olds’ categorization of animals than that of artifacts, the opposite was the case for both age groups when the information provided regarded the intent of the entities’ creator.

The above pattern is consistent with claims made by advocates of a domain-specific account of categorization. Backed by numerous studies (e.g., Diesendruck et al., 1998; Gottfried & Gelman, 2005; Jinson & Gelman, 2007; Rhodes & Gelman, 2009; Simons & Keil, 1995), researchers have argued that whereas children construe animals as natural kinds defined by intrinsic essential properties (Gelman, 2003), children view artifacts as defined by functional or intentional attributes, conventionally established (Bloom, 1996).

Importantly, the present experimental design allowed us to draw a more nuanced account of the interaction between these different types of conceptual information and perceptual similarity in determining children’s categorization decisions. Namely, as seen in Study 1, information about the internal properties of animals had such a dramatic effect on 5-year-olds’ categorization decisions that they almost completely ignored perceptual similarity (see Figure 3). In turn, whereas intentional information in Study 2 had a significant effect on children’s categorization of artifacts, its power to drive children to overlook perceptual dissimilarity was much milder than the one just described regarding internal information about animals (see Figure 4). For instance, 5-year-olds in the no-information conditions about animals in Study 1 and about artifacts in Study 2 rarely categorized test stimuli at the 12% similarity ratio as belonging to Category A. In contrast, 5-year-olds in the internal information condition about animals in Study 1 did so 75% of the time, compared with 5-year-olds in the intentional information condition about artifacts in Study 2 who did so about 13% of the time. This pattern is consistent with the notion that animal categories are defined by essences—for which internal properties work as proxies—whereas animal membership (Gelman, 2003). In turn, artifact categories seem to be less rigidly defined, with various factors, such as information about creators’ intent (Bloom, 1996), as well as perceptual similarity (Smith, 2005), functional information (Keil, 1989), conventional uses (Siegel & Callanan, 2007), and others, all contributing to children’s categorization decisions.

As noted above, the interaction between domain and type of conceptual property was especially salient among 5-year-olds. This was due to the finding that whereas 3-year-olds manifested differential responses to the two domains only in terms of the effectiveness of intentional information, 5-year-olds did so in regards to both, intentional and internal information. In other words, sensitivity to the domain specificity of internal information was relatively protracted. In general, this pattern seems consistent with numerous studies on children’s understanding of categories in both domains (e.g., Asher & Kemler Nelson, 2008; Gottfried & Gelman, 2005; Sobel et al., 2007). The contribution of the present findings lies in the fact that the pattern was manifested with the same stimuli and experimental design. It seems that whereas children view intentional information as exclusively pertinent to artifacts, they take internal information as potentially relevant across domains. Perhaps it is only with a developing understanding of biology that children come to realize the unique role of internal properties in the constitution of living kinds (Keil, 1995).

We investigated in Study 3 a relatively unexplored potential difference between domains, namely, whether labels carry differential weight in children’s categorization of animals and artifacts. The findings indicate that they do: 5-year-olds were more swayed by labels when categorizing artifacts than when categorizing animals.

The finding that labeling only had a marginal effect on 5-year-olds’ categorization of animals was somewhat surprising, as it seems to contradict previous studies (Gelman & Markman, 1986; Jaswal & Markman, 2007; Sloutsky & Fisher, 2004). A possible reconciliation between these findings is that whereas all these previous studies used relatively familiar animal categories as stimuli, the present study used novel animal categories. As Davidson and Gelman (1990) found, children were incapable of relying on animal categories defined by novel labels when the category members were too perceptually distinct. However, using a familiar category or a familiar name facilitated the use of labels in children’s inferences because it argued strengthened the perceived coherence of the category. Further consistent with the present findings, in a recent study using novel animated creatures as stimuli, Deng and Sloutsky (2012) found that 4- to 5-year-olds preferred categorizing based on a salient visual feature than on a label. Moreover, Noles and Gelman (2012) found that under optimal testing conditions, labels did not substantially alter children’s assessment of the visual similarity among potential animal category exemplars.

That labels affect artifact categorization is well documented (Gopnik & Sobel, 2000; Nazi & Gopnik, 2001; Sloutsky & Fisher, 2004, 2011). What the present findings add to this literature is that labels have a stronger effect on children’s categorization of artifacts than of animals. This finding suggests that contrary to domain-general accounts, labels do not have a uniform impact on children’s categorization. Rather, consistent with domain-specific accounts, children’s decisions on how to weight information relevant to categorization are informed by their naïve theories of the domains (see also Brandon & Gelman, 2009; Waxman & Gelman, 2010).

One domain-specific account of this differential effect of labels on children’s categorization again refers back to the notion of animal categories as defined by intrinsic essences causally responsible for animals’ physical appearances and behaviors (Atran, 1990; Gelman, 2003). Under this notion, labels serve as indices of these essences only as long as they do not blatantly violate how we perceive the natural groupings (Diesendruck, 2003). In turn, artifacts are by definition “relational”; that is, they are defined by how people relate to them, either via intentions, uses, conventions, or labels (Bloom, 1996; Keil, 1995). As such, labeling can more liberally mold artifact categories (see also Rhodes & Gelman, 2009).

The general picture deriving from all three studies is one that resonates with a distinction made by Locke (1707/1961): Animal categories—as natural kinds—are arguably believed to be defined by real essences, that is, an internal constitution from which all of the perceived qualities of a thing causally derive. Consequently,
the most reliable evidence for animal categorization is information about animals’ internal constitution. Artifact categories, in turn, can be said to be defined by nominal essences; that is, they are abstract humanly constructed ideas for which names stand (see also Schwartz, 1978). Consequently, the most reliable evidence for artifact categorization is information about how humans construe and name artifacts.

In this vein, the present findings show that 5-year-olds conceptualize objects from different domains in this Lockean manner. A critical question these accumulated findings invite is why children hold these different ontological commitments (see Sloman, Lombrozo, & Malt, 2007, for a general review)? How come children believe that animal categories are defined by internal properties somewhat immune to cultural factors, and vice versa for artifact categories? Is this differentiation a priori, as some evolutionary arguments might suggest (Atran, 1990; Csibra & Gergely, 2011), or a posteriori, based on children’s exposure to clustered patterns in the environment (Quinn, 2011; Sloutsky & Fisher, 2011) and culture- and language-specific construals (Malt, 1995)? We believe studies on infants and across cultures provide an excellent avenue to answer these questions.

References


Received July 30, 2012
Revision received December 10, 2012
Accepted December 26, 2012