Research Article

When Looks Are Everything
Appearance Similarity Versus Kind Information in Early Induction

Vladimir M. Sloutsky,1 Heidi Kloos,2 and Anna V. Fisher3

1Center for Cognitive Science, The Ohio State University; 2Department of Psychology, University of Cincinnati; and 3Department of Psychology, Carnegie Mellon University

ABSTRACT—The goal of this research was to examine mechanisms underlying early induction—specifically, the relation between induction and categorization. Some researchers argue that even early in development, induction is based on category-membership information, whereas others argue that early induction is based primarily on similarity. Children 4 and 5 years of age participated in two types of tasks: categorization and induction. Both tasks were performed with artificial animal-like categories in which appearance was pitted against category membership. Although the children readily acquired category-membership information and subsequently used this information in categorization tasks, they ignored category membership during the induction task, relying instead on the appearance of items. These results support the idea that early in development, induction is similarity based.

Inductive generalization is a critical aspect of human cognition because it enables people to generate new knowledge. For example, upon learning that a cat has a particular biological property, one can generalize this property to another cat. It is well established that induction appears early in development (Gelman & Markman, 1986; Mandler & McDonough, 1996; Sloutsky & Fisher, 2004a; Welder & Graham, 2001); however, mechanisms of early induction remain unclear. In an attempt to understand early induction, researchers have formulated two theoretical proposals, the knowledge-based approach and the similarity-based approach.

KNOCKLEDGE-BASED APPROACH

According to the knowledge-based approach (often referred to as the “naive-theory” position), when the task is to generalize properties of some natural-kind categories (such as animal kinds), induction is driven by conceptual knowledge. This knowledge is implemented as a set of conceptual assumptions that do not stem from parental input (Gelman, Coley, Rosengren, Hartman, & Pappas, 1998; see also Gelman, 2004, and Murphy, 2002, for reviews). First, young children are said to believe that individuals belong to more general categories, with members of the same natural-kind category sharing many important properties (i.e., the category assumption). Second, young children are said to believe that count nouns denote categories (i.e., the linguistic assumption). Although it is not claimed that the category and linguistic assumptions are a part of children’s explicit knowledge, it is generally argued that early induction is based on these assumptions. Specifically, proponents of the knowledge-based approach claim that when people, including young children, perform an inductive generalization, they first identify the category of an entity and then generalize properties of that entity to other members of the category. Therefore, even early in development, induction is said to be based on prior categorization of presented entities, and thus to be category based.

The main support for the category and linguistic assumptions comes from innovative research by Gelman and Markman (1986). In a series of experiments, they presented young children with a triad task, in which stimuli consisted of one target and two test items. The triad task was designed to pit appearance similarity against category membership: One test item belonged to the same category as the target but looked dissimilar to it, whereas the other test item looked similar to the target but belonged to a different category. Subjects were presented with a triad and were informed that one test item had a particular hidden property (e.g., “hollow bones”), and the other test item had a different hidden property (e.g., “solid bones”). The task...
was to generalize a hidden property to the target. Category membership was communicated by using the same label for the target and the dissimilar test item. In general, children were more likely to generalize to the target the property of the test item that shared the target’s label than the property of the test item that shared the target’s appearance (but see Sloutsky & Fisher, 2004a, Experiment 4, for diverging evidence and counterarguments). This finding was interpreted as evidence that children’s induction is based on category-membership information.

**SIMILARITY-BASED APPROACH**

According to the similarity-based approach, conceptual knowledge (e.g., knowledge that members of the same category share important properties) is a product rather than a precondition of learning. Therefore, early in development, cognitive processes do not depend on top-down conceptual knowledge. Instead, they are grounded in powerful learning mechanisms, such as statistical and attentional learning (French, Mareschal, Mermillod, & Quinn, 2004; Mareschal, Quinn, & French, 2002; McClelland & Rogers, 2003; Sloutsky & Fisher, 2004a; Smith, 1989; Smith, Jones, & Landau, 1996). Taking this approach, we have recently proposed a similarity-based model of early generalization—SINC, which stands for “Similarity-Induction-Naming-Categorization” (Sloutsky & Fisher, 2004a; Sloutsky, Lo, & Fisher, 2001).

According to SINC, early in development, both induction and categorization are based on the overall similarity of compared entities, with labels being features of objects that contribute to their overall similarity, rather than symbols denoting category membership. Support for this claim comes from findings that young children, but not adults, perceive identically labeled entities as looking more alike than differently labeled entities (Sloutsky & Fisher, 2004a; Sloutsky & Lo, 1999). Furthermore, this contribution of labels to similarity seems to stem from auditory input overshadowing (i.e., attenuating processing of) corresponding visual input early in development (Napolitano & Sloutsky, 2004; Robinson & Sloutsky, 2004; Sloutsky & Napolitano, 2003).

**INDUCTION: CATEGORY BASED OR SIMILARITY BASED?**

The proponents of both positions expect labels to affect induction, but differ radically in their view of the mechanisms that drive these effects. According to the knowledge-based approach, labels affect induction because they denote category membership, and category information drives induction. According to the similarity-based approach, labels affect induction because they contribute to the perceived similarity of items, and similarity drives induction. Therefore, reliance on category labels in a triad induction task is insufficient for distinguishing between the two positions.

One way to distinguish between these positions is to give subjects direct access to category information by teaching them a new natural-kind category that has a clear category-identification rule. Once subjects learn the category, they could be presented with an induction task, in which category membership is pitted against appearance. If category-based induction is a default for natural kinds, then young children (who successfully learn the category) should assume that members of the same kind have much in common. As a result, when performing induction, they should rely on category membership and ignore appearance information. Conversely, if similarity-based induction is a default, then young children (even when they successfully learn the category) should rely on appearance information, while disregarding category-membership information.

The experiments presented here had the following overall structure. Subjects were first presented with a category-learning task during which they learned that artificial animal-like creatures belonged to two natural kinds: Some were nice, friendly pets, and some were wild, dangerous animals. The membership in a category could be detected by a rule, whereas appearances were not predictive of category membership. The children were then given a categorization task with items that differed from those used during training. Subjects readily acquired the categories and accurately sorted the items according to their kind information. Then subjects were presented with a triad induction task. Each triad consisted of a target and two test items, with one test item sharing the target’s category membership but not its appearance, and the other test item sharing the target’s appearance but not its category membership. In the induction task, subjects were familiarized with a quasi-biological property of the target and asked to generalize this property to one of the test items. Finally, subjects were given a final (i.e., postinduction) categorization task using the same items as the induction task.

Predictions of the two theoretical approaches were straightforward: If children perform category-based induction, they should overlook conflicting appearances and generalize to the target properties from an item that they know belongs to the same kind as the target. Conversely, if children perform similarity-based induction, they should generalize properties on the basis of common appearances, despite their knowledge of category membership.

**EXPERIMENT 1**

**Method**

**Subjects**

Sixteen 4- and 5-year-olds ($M = 61.2$ months, $SD = 2.9$ months; 9 girls and 7 boys) participated in the experiment. Four addi-
types of stimuli, those with appearance A1 and those with appearance of an entity. Attributes were conjoined to create two categories. Within each category, items had appearance A1 or A2, according to the number of fingers and number of buttons (with no other features present). Therefore, category membership did not correlate with appearance.

Materials

The materials were colorful drawings of buglike entities, created by combining the following six attributes: a body, a tail, antennae, wings, buttons, and fingers (see Fig. 1). Each of the six attributes varied on two dimensions (i.e., size and color, color and shape, or shape and size), with each dimension having three “levels” (e.g., for the size of the wing: 1 = short, 2 = medium, and 3 = long). The resulting 12 attributes represented the appearance of an entity. Attributes were conjoined to create two types of stimuli, those with appearance A1 and those with appearance A2. For A1 stimuli, 88% of dimensions belonged to Level 1, whereas for A2 stimuli, 88% of dimensions belonged to Level 3. The rest of the dimensions for both A1 and A2 stimuli belonged to Level 2.

Two additional attributes of the entities, the number of buttons and the number of fingers, were used to identify category membership. Each number could range from 1 to 6. Category membership could be determined on the basis of the relation between these two attributes: Members of one category (C1 stimuli) had more fingers than buttons, whereas members of the contrasting category (C2 stimuli) had fewer fingers than buttons. Note that the category-identification rule had only minimal perceptual basis (i.e., the children were required to compare buttons and fingers to determine category membership). The use of such a rule ensured that category-membership information was not correlated with appearance and thus could be pitted against appearance.

Overall, four types of items were created: A1C1 and A2C1 items (i.e., stimuli that were members of C1, with A1 and A2 appearance, respectively), as well as A1C2 and A2C2 items (i.e., stimuli that were members of C2, with A1 and A2 appearance, respectively). Figure 1 shows an example of each type of item.

Design and Procedure

The experiment was administered on a computer and controlled by SuperLab Pro 2.0 software. The children were tested in a quiet room in their preschool by hypothesis-blind experimenters. The cover story involved a child who wanted to get a nice and friendly pet. However, the pet store carried two kinds of animals, nice and friendly ziblets and wild and dangerous flurps. The task was to determine whether an animal was a ziblet or a flurp.

The procedure consisted of four phases: category learning, initial categorization, induction, and final categorization. During category learning, the children were given information about the categories. Specifically, they were told: “To tell if an animal is a ziblet or a flurp, you have to count the buttons and the fingers. Ziblets always have more fingers than buttons.” Two examples followed, each consisting of the correct combination of the number of fingers and number of buttons (with no other features present). Then the children were presented with eight training trials, in which they were asked to determine whether a creature was a ziblet or a flurp. After responding, they received corrective feedback and were reminded of the rule for determining category membership. Note that during training, subjects were presented with A1C1, A1C2, A2C1, and A2C2 items (two training trials per item), so that only the rule (but not appearance) was predictive of category membership.

During the initial categorization task, the children were presented with new A1C1, A1C2, A2C1, and A2C2 items (again, two trials per item, with a total of eight trials) and were asked to determine whether each item was a ziblet or a flurp. No feedback was given, and the experimenter did not repeat the rule to the children. To be included in the sample, children had to perform correctly on at least six of the eight trials (i.e., respond with at least 75% accuracy). Four subjects were excluded because they did not reach this criterion.

After the initial categorization task, the children were presented with an induction task. They were told: “The pet-store owner has a few questions for those who want to buy a pet. Can you help get those questions right?” On each trial, the children were shown a triad consisting of a target item and two test items located underneath the target (see Fig. 2a); neither the target nor the test items had been used in the category-learning or initial categorization tasks. For a subset of children, the target was an A1C1 item, and for the rest of the children, the target was an A2C2 item. The two test items were A1C2 and A2C1 items (displayed next to each other, with their left-right position counterbalanced). On each trial, the children were told about a hidden property of the target (e.g., “thick blood”) and asked to pick the test item that had the same hidden property. Twelve induction trials were presented in random order, with each hidden property being paired with each test item a total of four times.

Fig. 1. Examples of the stimuli in Experiment 1. A = appearance; C = category. Items in Category 1 (C1), ziblets, were identified by having more fingers than buttons. All other items belonged to Category 2 (C2), flurps. Within each category, items had appearance A1 or A2, according to the values of their six attributes (body, tail, antennae, wings, buttons, and fingers). Thus, category membership did not correlate with appearance. Stimulus a is an A1C1 item, stimulus b is an A1C2 item, stimulus c is an A2C1 item, and stimulus d is an A2C2 item.
property being used twice. No labels were given during induction, with all items being referred to as “this one.” Notice that the selection of $A_1C_1$ and $A_2C_2$ items as targets and of $A_1C_2$ and $A_2C_1$ items as test items enabled us to directly pit appearance against category membership.

After the induction task, subjects were presented with a final categorization task, which was similar to the initial categorization task. We included the final categorization task to ascertain that the children could correctly categorize the induction items and that they did not forget the category-identification rule in the course of the experiment. The items used in the final categorization task were the test items used in the induction task intermixed with four catch items (the catch items were cartoonlike drawings of completely new animals and were used to control for overall alertness). The children were asked to determine whether or not each presented item was a ziblet. The catch items were correctly rejected by all subjects.

**Results and Discussion**

Results are presented in Figure 3. Subjects were highly accurate in the initial categorization task, $M = .95$, above chance, one-sample $t(15) = 27.8$, $p_{rep} = .998$, $d = 6.9$, and in the final categorization task, $M = .82$, above chance, one-sample $t(15) = 6.95$, $p_{rep} = .998$, $d = 1.74$. However, they ignored category information in the induction task, $M = .27$, below chance, one-sample $t(15) = -4.28$, $p = .001$, $p_{rep} = .99$, $d = -1.07$, relying instead on appearance information. This tendency to ignore category information was especially striking given that the children clearly knew which kinds the items belonged to: They accurately categorized items in the final categorization task, which used the same stimuli as the induction task.

Therefore, knowing which categories items belonged to, the children ignored this information when performing induction.

Could it be that the young children failed to perform category-based induction under these conditions because they could not simultaneously pay attention to the rule and generalize a hidden property? To eliminate this possibility, we conducted a control experiment in which a separate group of sixteen 4- and 5-year-olds participated in the same induction task. However, prior to the induction task, they were trained to rely on the rule when performing induction. Specifically, they were first told that “all animals that have more fingers than buttons also have the same stuff inside.” They were then given eight induction-training trials accompanied by corrective feedback. Finally, they were tested on 12 no-feedback induction trials using pictures of creatures and to-be-generalized properties different from those used in the training trials. Neither labels nor category information were given during induction training or testing. The results of the test trials indicated that the subjects reliably used the rule when inducing hidden properties, $M = .92$, above chance, one-sample $t(15) = 15.34$, $p_{rep} = .998$, $d = 4.34$.  

---

**Fig. 2.** Examples of triads used in the induction task of Experiment 1 (a) and Experiment 2 (b). In both triads, Test 2 belongs to the same category as the target.

**Fig. 3.** Proportion of category-based responses in the categorization and induction tasks of Experiments 1 and 2. Error bars represent standard errors.
Therefore, the results of Experiment 1 are unlikely to stem from young children being unable to perform the induction task.

It could also be argued that the results stemmed from the inclusion rule (i.e., ziblets have more fingers than buttons) being accidental rather than biologically important. To eliminate this possibility, we conducted a second control experiment with a separate group of sixteen 4- and 5-year-olds. The procedure was identical to the one used in Experiment 1, with the only difference being that there was an initial training phase in which the children were taught about the biological relevance of the category-inclusion rule. Specifically, they were told that ziblets catch their food with fingers (they have a chemical in their blood that makes their fingers really sticky), but do not need their buttons for anything. Despite the fact that the children accurately remembered the explanation at the end of the experiment, the results replicated those of Experiment 1. The children were accurate on both the initial and the final categorization tasks, Ms > .81, above chance, one-sample t(15) > 7.8, p_{rep} = .998, d_s > 1.96, yet they again ignored category information in the induction task, M = .31, below chance, one-sample t(15) = −5.56, p_{rep} = .998, d = −1.39, relying instead on appearance similarity.

Overall, the results of Experiment 1 point to an important dissociation: When appearance information is pitted against category information, children spontaneously use appearance and not category information to perform induction. These findings support the idea that early induction is similarity based, rather than category based.

EXPERIMENT 2

The goal of Experiment 2 was to test the boundary conditions of children’s similarity-based induction. One could argue that children induce properties on the basis of appearance similarity only when category information is in direct conflict with appearance. In other words, children might rely on similarity only when appearance is much more salient than category membership. At the same time, when appearances are not informative (i.e., all items look similar), children might rely on category membership, thereby exhibiting category-based induction. Experiment 2 examined this possibility by using an induction task in which the two test items were equally similar to the target, thus rendering appearances noninformative (see Fig. 2b). Experiment 2 eliminated the conflict between appearance and category membership, so that category membership was the only possible basis for induction. Because appearance information was not informative, an attempt to use this information when performing induction should have resulted in chance performance. Conversely, an attempt to use category information when performing induction should have resulted in below-chance reliance on appearance.

Method

Subjects
Fifteen 4- and 5-year-olds (M = 60.27 months, SD = 3.4 months; 8 girls and 7 boys) participated in the experiment. One additional child was excluded for not meeting the learning criterion (the learning criterion was the same as in Experiment 1).

Materials, Design, and Procedure
The stimuli and procedure were similar to those used in Experiment 1 with one critical difference: During the induction and final categorization tasks, a new set of test items was used, such that target and test items looked alike and one test item shared category membership with the target. As in Experiment 1, the final categorization task used items that were identical to the test items in the induction task, plus catch items.

Results and Discussion
Results are shown in Figure 3. As in Experiment 1, subjects were highly accurate in the initial categorization task, M = .98, above chance, one-sample t(14) = 26.25, p_{rep} = .998, d = 6.78, and in the final categorization task, M = .39, above chance, one-sample t(14) = 12.94, p = .0001, p_{rep} = .998, d = 3.34; their performance in the induction task did not exceed chance (M = .53, t < 1). Moreover, as indicated by the binomial probability test, none of the children exhibited reliable above-chance category-based induction performance. Therefore, even though category membership was the only predictive information, children did not use this information when performing induction. Instead, they may have attempted to use appearance information, which resulted in chance performance, given that this information was not predictive. These findings, together with the results of Experiment 1, indicate that young children do not spontaneously use category information when performing induction. Their spontaneous induction is based instead (when possible) on appearance information.

GENERAL DISCUSSION

The main finding of these two experiments is that young children induced hidden properties on the basis of appearance similarity rather than on the basis of shared membership in a natural-kind category. Furthermore, even when the category information was the only basis for induction, young children still did not use this information, although they could ably categorize the induction items. Although young children ably learned categories and used category information when categorizing items, they ignored this information when performing induction, relying instead on appearance similarity. This is direct evidence that early induction is similarity based and not category based. The results also indicate a dissociation between categorization and induction when the natural-kind category is not based on similar appearance.
A proponent of the knowledge-based approach could argue, of course, that the reported findings pertain only to unfamiliar categories, whereas category-based induction is limited to familiar categories. However, this argument has important theoretical consequences. If children’s conceptual assumptions are limited to familiar categories, these assumptions must be a product of learning and therefore cannot explain acquisition of new categories. Instead, the emergence of this conceptual knowledge itself requires an explanation. Furthermore, there is no reason why this conceptual knowledge could not be acquired by means of bottom-up associative learning. At the same time, if these conceptual assumptions are to provide a top-down non-associationist account of acquisition of new categories (as argued by Keil, Smith, Simons, & Levin, 1998, in their criticism of conceptual empiricism, and also by Gelman, 2004), they cannot be limited to familiar categories. Although results of the present study are consistent with the possibility that conceptual assumptions are the product of learning, they do challenge the possibility that conceptual assumptions are a priori.

A proponent of the knowledge-based approach could also argue that the reported results are limited to artificial situations in which there is no correlation between appearance and category membership. In our view, situations in which appearance and category membership do not correlate are not artificial—they are isomorphic to situations in which people have to decide, for example, whether a dolphin has a heart similar to that of a shark (i.e., same-appearance item) or to that of a bear (i.e., same-category item). Although situations in which appearance does not correlate with category information might not be very frequent, they are highly diagnostic as to the nature of induction. Note that under more regular conditions, appearance information and category information are highly correlated, so that it is difficult to distinguish between similarity-based and category-based induction.

The idea of pitting category information against appearance information was exploited by researchers arguing for category-based induction in young children (e.g., Gelman & Markman, 1986). However, in most of the studies using this idea, category information was conflated with linguistic information. For example, a child might be presented with a bird and told that the “bird” had a particular property, and then might be asked whether a dissimilar-looking “bird” or a similar-looking “bat” was more likely to have the same property. Induction from one bird to another bird has been often taken as evidence for reliance on category information. However, it is also possible that subjects in these studies relied on linguistic information, thereby exhibiting label-based rather than category-based induction.

The research reported here offers a novel paradigm that enables researchers to distinguish between category-based and label-based induction. Indeed, in both experiments, young children were trained to categorize entities into two natural-kind categories, and yet this kind of information played little or no role in the children’s induction. These findings seem to suggest that the reliance on category labels in the course of induction does not necessarily indicate that the induction is category based.

The empirical case for label-based induction is noncontroversial (e.g., Gelman & Markman, 1986; Sloutsky & Fisher, 2004a; Welder & Graham, 2001), although the mechanism underlying the effects of labels is hotly debated (see Sloutsky & Napolitano, 2003, for a summary of existing theoretical positions). At the same time, the empirical case for category-based induction early in development is much more controversial. Whereas some researchers have presented supportive evidence (see Gelman, 1988; Gelman & Markman, 1986, Experiment 3), others have presented evidence challenging the existence of category-based induction early in development (see Fisher & Sloutsky, 2005; Sloutsky & Fisher, 2004b); in both cases, the evidence has been indirect. The main contribution of the present study is that it provides direct evidence that when category information is pitted against appearance information and no labels are provided, early induction is based on appearance and not on category information. This research supports the idea that similarity-based induction is a default early in development, while challenging the idea of spontaneous category-based induction.

Acknowledgments—This research was supported by grants from the National Science Foundation (REC 0208103 and BCS 0078945) to V.M.S. We thank Chris Robinson, Nora Newcomb, James Cutting, and three anonymous reviewers for helpful comments.

REFERENCES


(Received 3/6/06; Revision accepted 4/17/06; Final materials received 5/5/06)