Categories and induction in young children

SUSAN A. GELMAN

University of Michigan

ELLEN M. MARKMAN

Stanford University

Abstract

One of the primary functions of natural kind terms (e.g., tiger, gold) is to support inductive inferences. People expect members of such categories to share important, unforeseen properties, such as internal organs and genetic structure. Moreover, inductions can be made without perceptual support: even when an object does not look much like other members of its category, and even when a property is unobservable. The present work addresses how expectations about natural kinds originate. Young children, with their usual reliance on perceptual appearances and only rudimentary scientific knowledge, might not induce new information within natural kind categories. To test this possibility, category membership was pitted against perceptual similarity in an induction task. For example, children had to decide whether a shark is more likely to breathe as a tropical fish does because both are fish, or as a dolphin does because they look alike. By at least age 4, children can use categories to support inductive inferences even when category membership conflicts with appearances. Moreover, these young children have partially separated out properties that support induction within a category (e.g., means of breathing) from those that are in fact determined by perceptual appearances (such as weight). Since

we examined only natural kind categories, we do not know to what extent children have differentiated natural kinds from other sorts of categories. Children may start out assuming that categories named by language have the structure of natural kinds and with development refine these expectations.

Whales aren't fish, fool's gold isn't gold, and glass-snakes are actually lizards. These anomalies pose a problem for theories of categorization: To any casual observer, whales look like fish, fool's gold looks like gold, and glass-snakes seem indistinguishable from snakes. As these cases illustrate, certain categories capture more than obvious perceptual features. Fish, gold, and lizards are examples of natural kind categories, classes of objects and substances found in nature. Analyses of natural kinds (Kripke, 1971, 1972; Mill, 1843; Putnam, 1970) point out some interesting properties of their category structure.

Mill (1843) noted that certain "kinds" are distinguished by the remarkable richness of their correlated features. These categories pick up dense clusters of information that extend far beyond our original characterization. For example, giraffes share a particular diet, life expectancy, gestation period, DNA structure, and so forth—attributes that are impossible to know by casual inspection. These properties are in addition to more obvious ones, such as four legs, long neck, spots, horns, and other perceptual characteristics. Many characteristics of giraffes are not known to the average person and some have yet to be discovered even by expert biologists. Sciences such as biology, physiology, and chemistry arose to study natural kinds, in part because there is a seemingly inexhaustible set of facts to be discovered and understood about these categories.

Mill's observations are supported by recent work on natural kind terms (Kripke, 1971, 1972; Putnam, 1970; Schwartz, 1977, 1979). As part of more extensive analyses, Kripke and Putnam point out that natural kinds cannot be characterized by simple perceptual properties. Although features may be set forth to identify members of a natural kind category, they often do not serve as necessary and sufficient criteria. For example, whales are shaped like fish and live and swim in water as fish do, but they are not fish. Conversely, whales are mammals even though they are not furry and do not walk on land, as other mammals do. Usually, of course, perceptual features reliably identify the category of an object. Giraffes, for example, have many perceptual characteristics in common. Yet attributes common to members of a category are often impossible to know by casual inspection, as Mill has demonstrated.

In summary, there are two points concerning natural kinds we want to emphasize. First, natural kinds have rich correlated structures, and second, they are not necessarily identified by simple perceptual properties. Both of

these points suggest hypotheses about how categories are used to support induction. The highly correlated structure of natural kinds suggests that new features learned about one category member will often be projected onto other category members as well. In this way natural kind categories should promote inductive inferences. Moreover, given that perceptual features do not always identify category members, there are two ways in which inductions could be made without perceptual support. First, even if an object does not look much like other members of a given natural kind, people should assume that it will share relevant properties with other members of the category. Second, these properties, such as internal organs or chemical structure, are often unobservable by the average person.

Clearly, only some sorts of inferences within a kind are justified. Among animal categories, for example, we expect members of the same species to share methods of reproduction, respiration, and locomotion. Yet we do not expect other sorts of properties (such as age) to be common even to members of the same species. Thus, at least implicitly, people have embedded natural kind categories in scientific or prescientific theories that limit what classes of properties are expected to be common to a given natural kind category.

Furthermore, not all terms name natural kinds. Some words (such as "square") are defined by necessary and sufficient features (such as "equilateral rectangle"). Others, for example words for simple artifacts such as furniture or clothing, may not have clear defining criteria, yet differ from natural kinds in that we do not expect to discover unlimited new facts about them. It is doubtful that we would have a science of chairs or of socks, in which new discoveries were made. Yet, we do have sciences devoted to individual animal and plant species because there are so many properties, as yet unknown, that are common to members of each of these types of categories.

Little is known about how expectations about natural kinds originate. How much exploration of categories or even explicit scientific training is needed before children come to expect that categories reflect more than superficial perceptual similarities? There is a large developmental literature suggesting that young children rely on superficial perceptual properties on Piagetian and other cognitive tasks (cf. Flavell, 1963, 1985). Young children have often been characterized as "concrete" and as "perceptually bound", meaning that their thinking is captured by the appearances of things. Tversky (1985) has found, for example, that a 4-year-old typically groups a fire engine with an apple because both are red, rather than grouping a fire engine with a car because both are vehicles. In this task and others, children seem unable to override what are sometimes misleading perceptual cues.

Based on such findings, one might expect that young children would rely heavily on perceptual characteristics of objects for judgments of category

membership. Young children may have no means of appreciating the rationale for grouping perceptually dissimilar objects together. Even for natural kinds, children might represent category members as sharing just superficial properties and only later come to realize that they also have deeper properties in common. Thus according to this view, children, with their reliance on perceptual features and their limited scientific knowledge, should not rely on categories to support inductive inferences about objects.

On the other hand, children might early on appreciate the highly correlated structure of at least some categories, despite their limited scientific knowledge. Even with only rudimentary scientific beliefs, children could notice that some categories have many observable features in common and then extend this belief, expecting natural kinds to be united by many unobservable properties as well. An appreciation for natural kinds at this early age would probably reflect an unsophisticated, undifferentiated belief in the richness of categories. Children would lack the requisite scientific knowledge that could limit their inductions to categories and attributes that are appropriate. Children may even fail to differentiate between natural kinds and other categories, at an early age.

In this work, we examined children's understanding of natural kinds. We investigated whether preschool children show an appreciation for the richness of category structure. As a first step, we tested children on only the clear cases: categories that, for adults, capture rich sets of nonobvious properties. Because we did not compare natural kinds and other sorts of categories (but see Gelman, 1984), the present work cannot determine whether natural kinds have any special status for young children. In other words, we will not be able to determine whether children are basing their inferences on the natural kind status of categories per se. We can, however, determine to what extent children rely on category membership to guide their inferences, when such inferences are in fact warranted.

One piece of evidence that children are not solely dependent on perceptual similarity for drawing inferences comes from work by Carey (1985). She showed several groups of children between ages 4 and 10 a mechanical monkey, one that could move its arms to bang cymbals together. Children, who knew that real monkeys breathe, eat, and have baby monkeys, were asked whether the mechanical monkey could breathe, eat, and have babies too. All but one group of 4-year-old children denied that the mechanical monkey possessed these animate properties. Despite the striking perceptual similarity of these two types of objects—mechanical and real monkeys—children did not generalize facts about one to the other. These children have differentiated living things from nonliving things and therefore refused to impute properties that characterize living things to nonliving things. At a very general level,

then — at the level of basic ontological distinctions (Keil, 1979) such as living versus nonliving — children do not rely solely on perceptual similarity for drawing inductive inferences.

In the present set of studies we questioned children about much more specific natural kind categories, from both biological (e.g., squirrel and snake) and nonbiological (e.g., gold and salt) domains. To determine whether children would induce new information from natural kind categories, rather than from perceptual appearances, we pitted category membership against perceptual similarity. Children were shown two objects and told a new fact about each. They then had to infer which of the facts applied to a third object that looked very much like one of the first two objects but was given the same category label as the other one. These experiments test whether very young children are sensitive to the highly correlated structure of natural kind categories and whether they use these categories, in the absence of perceptual support, to justify inductive inferences.

Study 1

Preliminary study with adults

One of the main assumptions about natural kind categories that motivates our developmental work is that adults will rely on the natural kind membership of an object more than its superficial perceptual appearance to draw inferences about internal structure, behavior, and other theoretically relevant properties. We tested this assumption in a preliminary study with 20 Stanford undergraduates.

In this preliminary study, twenty sets of picture cards were prepared with three pictures per set. Each triad was constructed so that the third picture looked like one of the first two pictures, but was from the same category as the other. A sample triad is illustrated in Figure 1. New information was given about the first two pictures, then a question was asked about the third picture. For example, on one problem adults saw a flamingo and a bat. Underneath the flamingo was written: "This bird's heart has a right aortic arch only." Underneath the bat was written: "This bat's heart has a left aortic arch only." Then below these two pictures was a picture of a blackbird (which looked more like the bat than the flamingo). Underneath the blackbird was written: "What does this bird's heart have?" After each choice, subjects were asked to rate their confidence in their answer on a scale from "1" (very unsure) to "7" (very sure).

Figure 1.   *Sample triad used in Studies 1–5.*
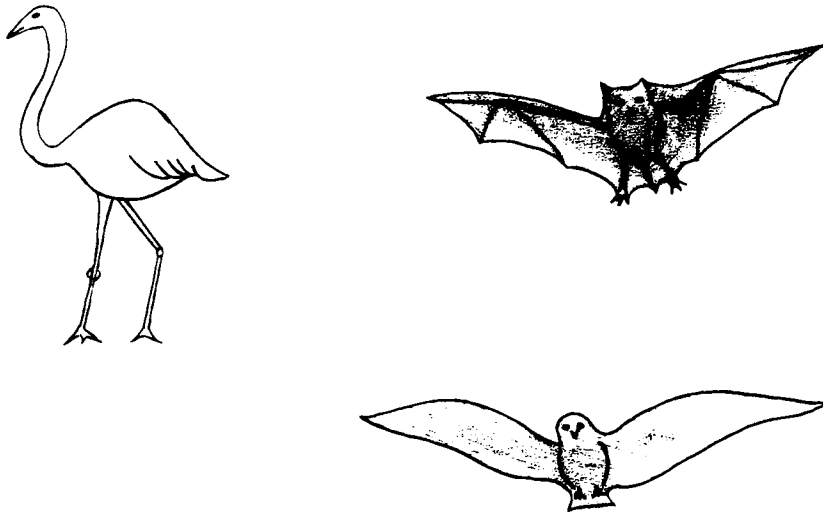
As predicted, adults based their inferences on the common natural kind membership of the objects, far more often than on their outward appearance. Overall, they concluded that the target picture had the same property as the other similarly labeled object an average of 86% of the time, which is significantly greater than chance, t (19) $=$ 15.77, $p <$ .001. Furthermore, subjects were highly confident that their choices were correct (mean rating was 5.8 on the 7 point scale, which is significantly greater than chance, t (19) $=$ 11.65, $p <$ .001).

Similarity ratings

It was important that the pictures that we thought were similar were in fact perceived as similar. Therefore we gathered ratings of perceptual similarity from adult subjects. Eighteen undergraduates participated to fulfill a course requirement for an introductory psychology class.

The 20 sets of pictures were divided into two pairs per set: The target picture paired with the picture that received the same category label and the target picture paired with the picture that was designed to be more similar in appearance. Subjects were given a set of written instructions, telling them that each pair should be rated on "how much alike the two pictures look to one another, on a scale of '1' (not at all similar) to '7' (extremely similar)." Subjects were told that they could take into account shape, color, size, com-

plexity, and individual features or parts. However, they were encouraged to make an intuitive judgment of overall similarity, rather than spending a great deal of time figuring out their answer. Examples of an "extremely similar" pair (two gray triangles) and a "not at all similar" pair (a small white square and a large black blob) were included.

Subjects' ratings of the pictures validated the experimental design of pitting perceptual similarity against category membership. Overall, the pictures that were designed to be more similar perceptually (e.g., blackbird and bat, or shark and dolphin) were rated as much more similar than the pictures that were from the same category (e.g., blackbird and flamingo, or tropical fish and shark). The similar pictures received a mean rating of 5.49, compared to a mean of 2.46 for the less similar pictures, min-F' $(1, 29) = 115.8, p < .001$.

Study with children

Young children might, like the adults, infer a new property of an object from its category, or they might instead be governed by the perceptual appearances of the objects. In this study, children were tested on the same categories that adults were questioned about. However, because they would not understand most of the attributes that adults had heard, children were asked about different ones.

Method

Subjects
Subjects included 59 children from three preschools in northern California, with 19 children in the experimental condition, and 20 children in each of the other two conditions. Subjects ranged in age from 4;0 to 4;11, with a mean age of 4;5.

Design
There were three conditions in this study. In the Experimental condition children were taught information about each of two objects. They were then shown a third object that looked like one of the two training objects but was given the same category label as the other. Children were asked to infer which information applied to the third object. This condition was designed to reveal whether children's inferences are influenced by their knowledge of an object's category or by the perceptual similarity.

A second condition, the No-Conflict Control, was designed to demonstrate that when perceptual similarity and category membership coincide, children readily draw the correct inferences. In this condition, like the first, children

were taught properties of two objects. But in this case, the third object not only looked like one of the training objects but was also given the same label as that object.

A final condition, the Attributes control, was designed to make certain that children did not already know the information we would be teaching them. Children saw only one picture at a time—the third item in the other two conditions—and were asked which of the two properties applied. Children were expected to perform at about chance level in this condition.

Procedure

Each child was tested individually, by one of two experimenters. To maintain children's interest throughout the 25 minutes that this task required, children were given a sticker after each set of pictures to place on a sheet that had blank balloons drawn on it.

Experimental condition.   Children saw 20 sets of three pictures each. Information was given concerning two of the pictures in each set, and children were asked a question about the third picture. The third picture looked like one of the first two pictures but was given the same category label as the other. For example, children were shown a tropical fish, a dolphin, and a shark. In this case, the shark was perceptually similar to the dolphin but was

Table 1.   *Sample items and attributes used in Studies 1 and 1a*[a]

Biological items

This squirrel eats bugs. (gray squirrel)
This rabbit eats grass. (brown rabbit)
(Target:) squirrel (kaibab, looks like rabbit)

This dinosaur has cold blood. (brontosauros)
This rhinoceros has warm blood. (gray rhinoceros)
(Target:) dinosaur (triceratops, looks like rhinoceros)

Nonbiological items

If you put this gold in a hot oven, it melts. (gold bar)
If you put this clay in a hot oven, it burns. (reddish blob)
(Target:) gold (brown blob; looks like clay)

This pearl comes from inside a sea animal. (seed pearl)
This marble comes from a big piece of rock. (round, pink)
(Target:) pearl (round, pink; looks like marble)

---

[a]Descriptions of the objects used are given in parentheses. These descriptions were *not* mentioned to subjects.

given the same label as the tropical fish. The experimenter labeled the three pictures, "fish" for the tropical fish, "dolphin" for the dolphin, and "fish" for the shark. Children were asked to repeat the names until they could name all three pictures correctly. (On 88% of the trials, children repeated all three names correctly on their first try.) The experimenter then pointed to the tropical fish and said, "This fish stays underwater to breathe." She pointed to the dolphin and said, "This dolphin pops above the water to breathe." She then pointed to the shark and said, "See this fish? Does it breathe underwater, like this fish, or does it pop above the water to breathe, like this dolphin?" Further example items are presented in Table 1.

No Conflict Control. The picture triads and procedure used in this condition were identical to that of the Experimental condition, except that the labels of the two similar objects were made to agree rather than conflict. To continue with the earlier example, in the No Conflict condition children heard the tropical fish labeled "fish," the dolphin labeled "dolphin," and the third picture (shark) labeled "dolphin" (instead of "fish" as in the previous condition). The children in this condition were provided with the same information about the two initial pictures, i.e., that the fish breathes underwater and the dolphin pops above the water to breathe. The experimenter then pointed to the third picture and said, "See this dolphin? Does it breathe underwater, like this fish, or does it pop above the water to breathe like this dolphin?" Children in this condition received only 18 questions because 2 of the 20 experimental items could not be plausibly relabeled.

Attributes Control. Children in this condition viewed only one picture at a time—the third item from each of the triads in the Experimental condition. Without hearing any prior information, children were asked to judge which of the two properties applied. For example, children would be shown the picture of the shark and asked, "See this fish? Does this fish breathe underwater or does it pop above the water to breathe?"

The items were presented in a different random order to each child, with the constraint that all items of a given type (biological or nonbiological) were given in a block, with order of the blocks counterbalanced across subjects. For each triad, the order in which the two training pictures were shown and attributes given was counterbalanced across subjects and across items. Subjects were randomly assigned to the different conditions, with the constraint that these conditions were balanced for age and sex.

Results and discussion

The main question that this study was designed to address is whether pre-school children are willing to infer properties of an object based on its natural kind category. In particular, when category membership and perceptual simi-larity are in conflict, will children show any sensitivity to the category mem-bership or will their inferences be based on the appearance of the objects? Table 2 presents the data according to condition and item type.

Table 2.   *Percent correct (category choices) in Study 1*

	Experimental condition	No Conflict Control	Attributes Control
Biological	67	89	59
Nonbiological	69	87	48

Analysis of choices

To address these questions, we first needed to establish that children's inferences in the Experimental condition were based on the information pro-vided to them in the experiment rather than on preexisting knowledge. The results of the Attributes Control condition indicate that children were in fact unaware of the correct answers. When simply given the test question with no extra information to guide their answer, children were performing at chance level, answering a mean of 53% of the questions correctly.

On the other hand, when the perceptual similarity and category label coin-cided, in the No Conflict Control, children were quite capable of drawing the correct inference. When both the label and the appearance of the object led to the same conclusion, children were correct on 88% of the items, which is significantly better than chance, t (19) = 10.77, $p < .001$.

In the Experimental condition, where perceptual similarity and category membership were opposed, children preferred to use the category informa-tion 68% of the time, which is significantly better than chance, t (18) = 3.78, $p < .001$. Thirty-seven percent of the children consistently based their judg-ments on common category membership. That is, they inferred the property of the new object based on category membership on at least 15 out of 20 items ($p < .05$ by a sign test). In contrast, no child showed a consistent preference for basing their inferences on the perceptual similarity of the objects.

A 4-way ANOVA was performed, with condition (experimental, no-con-

flict, attributes control), sex, and order of presentation (biological versus nonbiological items first) as between-subject variables, and category type (biological versus nonbiological) as a within-subject variable. There were three significant effects.

First, there was a main effect of condition, F (2,47) = 21.62, p < .001. Each condition differed significantly from the other two, by separate t-tests (p < .01). Thus, the children in the Experimental condition were taking account of the training information in deciding about the properties of the new objects, as performance in that condition was better than in the Attributes Control condition. Despite children's high performance in the Experimental condition, they were still somewhat misled by perceptual appearances as shown by the 20% decrement from the No Conflict Control condition. Importantly, however, adults showed a similar pattern, drawing inferences based on appearance 14% of the time when category and appearances conflicted (see Preliminary Study with Adults). Thus, 4-year-olds showed a remarkably adult-like pattern.

A min-F' analysis, that treats both items and subjects as random effects, was used with the Experimental group and Attributes Control group. Sex and category type were included as fixed effects. The No Conflict group was excluded because it had 18 rather than 20 items. The difference between these two conditions was significant, min-F' (1, 53) = 5.46, p < .05, indicating that the effect would generalize to new categories as well as subjects. Children were able to use category membership as often for the biological as for the nonbiological categories. Overall the mean correct was 67% for the biological and 69% for the nonbiological categories.

The second significant effect was a Category Type × Condition × Sex interaction, F (2,47) = 4.38, p < .02, which was marginally significant when tested by a min-F' statistic that included the Experimental and Attributes Control condition, min-F' (1,45) = 3.12, p < .10. This interaction was due primarily to a sex difference in the Experimental condition. Boys were above chance on the biological categories (73%), t (11) = 4.69, p < .001, but not on the nonbiological categories (66%), while girls showed the reverse pattern, answering above chance on the nonbiological categories (74%), t (6) = 2.71, p < .05, but not on the biological categories (57%). We have no explanation for this difference, although it is possible that boys and girls differ as to which categories they find more familiar. Boys may be more familiar with snakes, worms, bugs, leaves, etc.—that is, the biological categories we tested, whereas girls may be more familiar with sugar, salt, diamonds, pearls, etc.— the nonbiological categories we tested.

Finally, there was a Category Type × Presentation Order × Sex interaction, F (1,47) = 7.07, p < .01. Because this effect held for data collapsed

across conditions, it is difficult to interpret meaningfully. However, it seems largely carried by the Experimental condition, where the sex difference was most striking in the first block of items. Performance in the second block did not yield any differences between biological and nonbiological categories for either boys or girls.

Similarity ratings

To determine to what extent similarity accounted for children's performance, correlations were computed between children's accuracy in the Experimental condition and the similarity ratings of the adults. Both similarity ratings and percent of category choices across categories were converted to rank-order scores before correlations were computed. Table 3 presents the correlations between accuracy in selecting category members and (a) the similarity ratings of the category members, (b) the similarity ratings of the perceptually similar pairs, and (c) the difference score of (a) minus (b). In this study, none of the similarity ratings reliably predict how often children based their inferences on membership in a common category.

Table 3.   *Correlations between children's inferences (Studies 1, 1a, and 3) and adult similarity ratings (Spearman's rho)*

Number of category selections	Similarity ratings		
	Common category pairs	Perceptually similar pairs	Difference score
Study 1			
Biological	−.04	−.12	−.18
Nonbiological	.15	−.09	.26
Study 1a			
Biological	.24	−.21	.23
Nonbiological	.39	−.54	.53
Study 3			
Biological	.43	−.53	.65[a]
Nonbiological	−.43	.08	−.32
Expected direction	+	−	+

[a] $p < .05$.

Study 1a

Study 1a was a modified replication of the Experimental condition of Study 1. Informal observation in the first study suggested that the children may have worried about finding the correct answer. For this reason the instructions were modified to explicitly point out that there were two good answers and the child needed to pick the best one.

Eighteen 4-year-olds (mean age of 4;5) participated in the study, none of whom had participated in Study 1. The results of Study 1a replicated those of Study 1. Children believe that a category label is a better clue to an object's behavior and internal structure than is outward appearance. Overall, children drew inferences based on category membership an average of 73% of the time, which is better than chance, t (17) = 5.23, p < .001. Despite the compelling perceptual similarity of two of the objects in each triad, no child chose consistently (i.e., above chance) on the basis of appearance. In contrast, 50% of the children consistently chose on the basis of common category membership.

In summary, for many 4-year–olds, natural kind categories such as *squirrel* and *diamond* promote a rich set of inductive inferences. These young children have already come to expect new knowledge to be organized in accord with the categories named by their language even in the stringent test case where the label conflicts with outward appearances. These results are quite surprising in two respects: first, children typically rely on perceptual similarity on cognitive tasks. And second, research by Carey demonstrates that children lack the scientific knowledge that could justify their categorical choices here. We will return to this issue in Study 3.

Study 2

In Study 1, children relied more on category membership than perceptual similarity to draw inductive inferences. Identity of categories was always conveyed by identity of the category labels. One question this raises is whether identity of the linguistic information is necessary for children to use the common category as a basis for inductive inferences, or whether other means of indicating common category membership would be sufficient. To address this question, we ran a Synonyms condition where category membership was conveyed by means of synonyms rather than identical labels. If children infer that objects named by synonyms (e.g., "rabbit" and "bunny") share the same properties, then their within-category inferences cannot be based simply on identity of the labels. Unfortunately, however, there are not many synonyms

for natural kinds in the vocabulary of 4-year-olds. So to supplement the few synonyms we could find, we also included categories that were related by subset-superset relations, e.g., "rose-flower." These pairs in addition to having different labels, would also require children to make an additional inference (e.g., that a rose is a flower) to use the common category as a basis for drawing an inductive inference. Therefore, they should underestimate the categorical inferences children would make if genuine synonyms could be found.

Another question that Study 2 was designed to address is whether children would use the common category to make even arbitrary decisions for which common category membership is not relevant. To test this we ran an Arbitrary Decision condition where children were asked to decide what color marker should go on a picture after witnessing the experimenter place one colored marker on a perceptually similar picture and a different colored marker on a dissimilar picture with the same category label. If children are distinguishing between the induction task where category labels are relevant and this arbitrary task, where they are not, then children should be at chance in this condition, having no real basis on which to make a decision.

Both conditions just described provide strict tests of children's reliance on categories. Children may fail to use the appropriate information in each condition, even if they possess an adult-like understanding of category structure. In the Synonyms condition, children must make additional inferences (e.g., a rose is a flower; a bunny and a rabbit are the same sort of thing) in order to rely on the category. In the Arbitrary Decision condition, children have no rational basis on which to make a decision, and so may rely on the category labels by default. This would not be an *incorrect* decision, simply an unmotivated one.

Method

Subjects

Subjects included 48 preschoolers, none of whom had participated in Studies 1 and 1a, from six nursery schools in northern California. Children were tested in three conditions: the Synonyms condition, the Arbitrary Decision condition, and the Standard condition. The age range and mean age for each condition were: Synonyms condition (4;3 to 5;8, 4;10); Standard condition (4;0 to 5;6, 4;9); Arbitrary Decision condition (4;0 to 5;6, 4;9). Nineteen of the children were girls.

Procedure

The task was introduced to all children as a game, in which after every set

Table 4.   *Sample items and attributes used in Study 2*[a, b]

This puppy [baby dog] hides bones in the ground. (brown dachshund)
This fox hides food in the ground. (red fox)
(Target:) puppy (red dog, looks like the fox)

This snake [cobra] eats plants. (cobra)
This worm eats meat. (earthworm)
(Target:) snake (small brown snake, looks like worm)

When you put this sand [the desert] in water, it just gets wet. (hard lump of reddish sand)
When you put this sugar in water, it melts. (pile of brown sugar)
(Target:) sand (pile of brown sand, looks like the sugar)

This rock [stone] comes from a mountain. (rounded blob)
This chalk comes from the ocean. (angular, flat-surfaced)
(Target:) rock (angular, flat-surfaced; looks like the chalk)

---

[a]Alternative labels are given in square brackets.
[b]Descriptions of the objects used are given in parentheses. These descriptions were *not* mentioned to subjects.

of pictures the child would receive a sticker to place on a sheet of paper to make a picture. Before presenting the experimental items, the researcher first asked the child three easy warm-up questions (e.g., "Is this [green bow] green or is it red?").

Children in the *Synonyms* condition received eight problems, each consisting of a triplet of pictures constructed to be analogous to those in the earlier studies. Unlike the earlier studies, however, category membership was conveyed by means of synonyms. For example, one triplet consisted of a target rabbit, another rabbit with different appearance, and a squirrel that had long ears and looked like the target rabbit. The two rabbits in the Synonyms condition were called "rabbit" and "bunny," respectively. Examples of items and attributes can be found in Table 4.

Children in the *Standard* condition received the same eight sets of items as in the Synonyms condition, except that objects from the same category were given *identical* labels rather than synonyms. This task is a replication of Studies 1 and 2, except for the items used. So for example, on one item children saw two rabbits and a squirrel — and both rabbits were labeled "rabbit."

Children in the *Arbitrary Decision* condition saw the same eight sets of items and heard the same labels for the pictures as in the Standard condition. However, the experimenter presented the child with an arbitrary, game-like task rather than with a series of inferencing problems. That is, for each

problem the experimenter placed a small card with a colored dot on each of the first two pictures. For example, the experimenter placed a red dot on one rabbit and a yellow dot on the squirrel. The child then had to pick which color dot to put on the third picture (the other rabbit that looked like the squirrel): a yellow dot or a red dot.

Results and discussion

In the Standard condition, designed to replicate Study 1, children once again drew inferences to category members above-chance (68% of the time, t (15) = 2.83, $p < .05$). Moreover, as the results from the Synonyms condition indicate, children do not need to hear common labels to use common category membership to draw inferences. In the Synonyms condition, where children heard either synonyms or subclass-superclass relations to indicate common category membership, children still were basing their inferences on category membership greater than chance (63% of the time, t (15) = 3.87, $p < .005$).

The results of this study further indicate that children have begun to differentiate between inferences where category membership is relevant from arbitrary decisions where it is not. When children were asked about arbitrary decisions, e.g., what color chip should go on a given picture, they were not significantly more likely to base that decision on the color chip they had seen placed on a common category member than a chip placed on a perceptually similar picture. Children in the Arbitrary Decision condition were performing at chance levels (57% category choices, t (15) = 1.45, $p > .1$).

The results in the last two conditions were very close (63% versus 57%). It appears that children may have a slight bias toward basing their answers on category membership, even in the Arbitrary Decision task. Nonetheless, responses in the two conditions look clearly different when the patterns of individual children are considered: only 6% of the children in the Synonyms condition answered based on appearance more than half of the time, whereas 25% of the children in the Arbitrary Decision condition did so.

In summary, children's inferences were based on the category membership of the objects in question, and not simply on how the pictures were labeled. First, children drew inferences within each category even when category members were not given identical labels. And second, when category members *were* given identical labels, children relied on these category names above chance only when the task required them to draw inferences.

Study 3

The purpose of Study 3 is to examine whether preschool children are selective in the kinds of inferences that they make. Children this young have impoverished scientific knowledge (Carey, 1985; Dickinson, 1982), yet firmly believe that category members have more in common than meets the eye. This implies that preschoolers' beliefs about categories may be a fairly general one, unchecked by specific world knowledge. In other words, preschoolers may not have sorted out which properties legitimately do and do not promote inferences within natural kind categories. Instead, they might erroneously infer information on the basis of common category membership, even when asked about attributes that are more likely to be consequences of physical or perceptual properties.

The properties that were examined in the previous studies are ones that in fact should be largely determined by the natural kind category of the objects. There are cases, however, where inductive inferences based on category membership would be unwarranted. Whether a substance is solid or granular, for example, is a good predictor of whether it will blow away in the wind; its category membership is less predictive. Note that even in this simple example, the perceptual information alone is not sufficient to draw the conclusion. Some, at least primitive, knowledge of the relation between granularity and stability is needed. The answer is not given perceptually but must be derived from the perceptual information. If attention to category membership dominates children's judgments regardless of type of property, then children should assert, for example, that a rock-like chunk of salt will blow away in the wind as does fine-grained salt, and not remain in place as does a rock.

Method

Subjects
Forty children, 20 girls and 20 boys, from nursery schools in northern California participated in this study. The children ranged in age from 4;0 to 5;0, with a mean age of 4;6. None of these children had participated in the earlier studies.

Procedure
Children were randomly assigned to either the Experimental or the Control condition with the restriction that the two conditions were balanced for age and sex. Children in the Control condition were given the 20 test questions without any extra information. For example, children were shown the shark

Table 5.   *Sample items and attributes used in Study 3*[a]

Biological items

This bird's legs get cold at night. (flamingo)
This bat's legs stay warm at night. (black bat)
(Target:) bird (blackbird, looks like a bat)

This snake eats a cupful of food. (cobra)
This worm eats a spoonful of food. (earthworm)
(Target:) snake (brown snake, looks like worm)

Nonbiological items

If you squeeze this diamond, nothing happens. (round, uncut diamond)
If you squeeze this glass, it'll cut your hand. (cut glass)
(Target:) diamond (cut; looks like the glass)

This cotton is a soft place for a spider to sleep. (cotton ball)
This silk is too hard for a spider to sleep on. (swatch of cloth)
(Target:) cotton (swatch of cloth, looks like silk)

[a]Descriptions of the objects used are given in parentheses. These descriptions were *not* mentioned to subjects.

and heard, "This is a fish. Do you think it weighs 20 pounds or 100 pounds?"

The procedure for the Experimental condition was identical to that of Studies 1 and 2, except that different attributes were being taught. Children viewed three objects, two placed side-by-side and one centered beneath them. Each of the objects was labeled and children were asked to repeat the labels. A property was attributed to each of the two topmost pictures, and then children asked which of the properties applied to the third target picture. Examples of the items selected are presented in Table 5. For example, after the pictures were labeled, children were told of a tropical fish: "This fish weighs 20 pounds." They were told about an object from a different category, in this case a dolphin, that, "This dolphin weighs 100 pounds." They were then asked about the target picture, a shark that looked more like the dolphin than the tropical fish: "See this fish? Does it weigh 20 pounds like this fish or 100 pounds like this dolphin?" Children were asked to justify their choice after each item.

Results and discussion

In this study children were asked questions about perceptually based attributes: weight, visibility at night, and so forth. With no prior information on

Table 6. *Patterns of responses of subjects in Studies 1, 1a, 2, and 3*

	Number of categorical choices			
	0–5[a]	6–9	10–14	15–20[b]
Percent subjects				
Study 1	00	16	47	37
Study 1a	00	17	33	50
Study 3	20	35	25	20
	Number of categorical choices			
	0–1[a]	2–4	5–6	7–8[b]
Percent subjects, Study 2				
Standard	06	19	44	31
Synonyms	00	31	63	06

[a]below chance, $p < .05$.
[b]above chance, $p < .05$.

which to make inferences, children in the Control condition selected the category choice no more than expected by chance (48% of the time). In contrast to the earlier three studies, fewer children in the Experimental condition based their inferences on common category membership. They selected on the basis of common category an average of 49% of the time which is not significantly different from chance or from the control group. Yet the Experimental condition, unlike the Control, was markedly trimodal. Four of the 20 children consistently selected on the basis of perceptual similarity, that is, they chose the attribute of the perceptually similar picture for at least 15 out of 20 items, and four children consistently selected on the basis of common category membership even though that choice was unwarranted. The remaining 12 children did not seem to know which answer to choose, with a mean of 47% categorical choices. This condition contrasts with the Control condition, in which all 20 children fell into the "undecided" group. (See Table 6 for the results.)

From these data, it is clear that at least some children are sensitive to the type of attribute involved when drawing inferences. Children in this study relied less on the category than in Studies 1, 1a, and 2. Overall, children were sensitive to the differences between attributes such as weight, that are consequences of perceptual properties, and attributes such as means of breathing,

that are common to members of a species. Twenty percent of children in this study consistently based their judgments on perceptual similarity, whereas only 1% did so in the earlier three studies of attributes common to a natural kind category.

On the other hand, some children based their inferences predominantly on common category membership, thereby overgeneralizing the importance of the category label. It was unwarranted for children to assume, for example, that a legless lizard can run as quickly as a four-legged lizard, or that a large pearl weighs as much as a smaller pearl. These children understood the importance of category names to promote induction, but were not selective in their inferences. In part this result may be explained by a closer consideration of the attributes used in this study. Each was related only indirectly to a perceptually obvious feature. For example, one item asked how well a short-eared rabbit can hear. Unless a child realizes the link between ear-size and hearing ability, he or she will miss the relevance of appearance for this attribute. It could be that children are most sensitive to the distinction between "deep" and "shallow" attributes, when the perceptual basis of the attributes is especially clear.

We found some support for this hypothesis in a small supplementary study with adults. Eight graduate students (naive to the purpose of the study) were asked to rate how much each of the properties in Study 3 "depends on some *visually perceptible* property (such as size, color, or shape)," using a scale of 1 ("appearances are NOT relevant") to 7 ("appearances are VERY relevant"). We then correlated the ranks of these ratings, on an item-by-item basis, with the ranks of children's category choices in Study 3. We would expect the correlations to be negative, if children's responses reflect how "perceptual" each property seems to adults. The correlations were .13 and −.70, for biological and nonbiological categories, respectively. At least for the nonbiological categories, children are significantly more likely to draw inferences based on appearance when the property clearly calls for an appearance-based answer ($p < .05$).

Similarity ratings

Table 3 presents the Spearman's rho correlations between children's inferences in this study and adult similarity ratings, on an item-by-item basis. For biological categories, perceptual similarity of the pictures significantly correlates (on one of three measures) with children's performance on the inferencing task. That is, the perceptual pull of the similar picture predicts the proportion of choices based on appearance.

In sum, two factors affect children's performance in this study: (1) for biological categories, the perceptual pull of the similar picture; (2) for non-

biological categories, the perceptual relevance of the property. Thus, children are at least somewhat selective as to when they draw inductive inferences based on category membership.

Yet some children draw many more inferences based on category membership than others, on this task. The distribution of results in Study 3 is clearly non-random (fully 40% of the children performed significantly above or below chance). This age may be a transitional period: children may start out overgeneralizing the importance of the category label, then gradually refine their expectations, drawing inferences only for relevant properties such as internal structure (not weight or visibility).

General discussion

On each of a series of problems, children had to decide whether a given object possessed one or the other of two attributes. On hearing the bare question alone, with no prior information to guide them, children had no basis for forming an induction and, as expected, simply guessed at which property applied. In two other conditions, children were first told which attribute applied to each of two training objects. In the simple case where one of the training objects matched the target object in both appearance and category, 4–year–old children almost always drew the appropriate inference. For example, children would infer that a second dolphin (large and gray) pops out of the water to breathe, after finding out that a first dolphin (large and gray) pops out of the water to breathe. Their performance in the No Conflict condition was excellent, and establishes that the simple inferential problem is well within the capacity of young children.

The most informative condition was one in which perceptual appearance and the category of an object led to divergent conclusions. Children were taught a property of one object and a different property of a second one. They were then shown a third object that looked much like one of the first two but was given the same category label as the other. For example, children saw a tropical fish and were told that it was a *fish* and that it breathes underwater. They saw a dolphin and were told it was a *dolphin* and that it pops out of the water to breathe. They then had to decide how a second *fish*, a shark that looks like a dolphin, breathes. Children relied on the shared category to promote inductions even in this stringent case where perceptual similarity would lead to a different conclusion. Moreover, children's inferences are based on common category membership and not just on identity of labels. In Study 2, when members of the same category were given synonymous labels (e.g., "rock" and "stone") or subclass-superclass labels (e.g., "rose"

and "flower"), children still preferred to draw inferences within the category, above chance.

These results are at odds with a widely held view that children's thinking is strongly influenced by the perceptual appearances of things. Several of our findings suggest that children are not dominated by appearances either in their conception of the structure of categories or in their use of categories to support inductions. First, most children accepted our label for the third object even though it looked more like a member of the other category. For example, one object was a squirrel with very long, rabbit-like ears. Overall, it looked more like a typical rabbit than a squirrel. Some children noted the discrepancy, some even mildly objected to the label, but for the most part, children accepted the labels for these abnormal category members. This finding is consistent with recent work by Flavell, Flavell, and Green (1983) on the development of the appearance-reality distinction. Like our subjects, preschool children in their studies accepted category labels even in the face of discrepant appearances (e.g., a sponge that looks convincingly like a rock is called a "sponge" by these children).

Second, only one out of 69 children in the Experimental condition, in Studies 1, 1a, and 2, consistently generalized properties on the basis of perceptual similarity between objects. Even though the rabbit and the rabbit-eared squirrel looked very much alike, children did not assume that they both ate grass like the rabbit did. Rather, children reliably used the category of the object to support inductive inferences, even when this conflicted with the appearance of the objects.

The only kinds of properties asked about, in Studies 1 and 1a, were ones that were reasonable to project from one category member to another. We asked about the eating habits, means of breathing, and internal organs of biological categories, and about chemical and physical properties of non-biological categories. The results of Studies 2 and 3 indicate ways in which children have begun to limit the importance of the category. First, when the task involves only some arbitrary decision to be made, children are not biased to base the decision on category membership. Second, children have begun to limit induction proper. In Study 3, we asked children about properties that should generalize on the basis of perceptual properties rather than category membership. For example, children were shown a rock and told that it would not blow away in the wind, and some fine-grained salt and told that it would blow away in the wind. They were then shown a rock-like chunk of salt and asked whether this salt would or would not blow away in the wind. In answering such questions, a few children did consistently rely on category membership. These children have overgeneralized the relevance of the category for induction. Most children, however, did not reliably use the category to sup-

port the inductions. For example, they did not consistently infer that the rock-like chunk of salt would blow away in the wind as would the fine-grained salt. Across several studies, those children who were asked about perceptually based properties were the only ones to reliably use the perceptual appearances of objects to support their inductions.

By age 4, children use category membership to support inductions, even in the stringent test case where perceptual appearance and category label lead to different conclusions. Moreover, children have begun to differentiate between the kinds of properties that are justified to project to other category members and those that are not. Despite all these accomplishments, however, there is much left for children to learn about the relationship between categories and induction.

The first issue is how children constrain which kinds of categories support inductions. Some categories, such as artifacts (e.g., chairs, coats), do not pick out objects in nature that have indefinitely many properties in common. We assume that adults do not expect artifacts such as forks or saws to have as rich an internal structure as natural kinds. The present studies examined only natural kinds, and so do not tell us whether natural kinds have a special status for children. More recent work (Gelman, 1984) suggests that preschoolers do not yet consistently distinguish between natural kinds and artifacts in the inferences they draw.

A related issue is that children must sort out which properties are likely to be common to members of different types of categories. Although they have begun to work out this problem by age 4, their distinctions on even a crude level are imperfect. Even with properties that are determined by perceptual features (e.g., the weight of an object), some children based their inductions on the category. For example, there were children who claimed that a large fish weighs the same amount as a little fish, because both are fish. Furthermore, this problem quickly becomes more complicated. Adults' inductive inferences are constrained in ways that no one has yet been able to characterize (Goodman, 1955); we lack good theories to explain which inferences are warranted. A predicate may or may not promote inductions, depending on the level of abstraction of a category (all *dogs* bark, but not all *animals* bark) or the scientific domain (density at room temperature is important for metals but not animals). There are other principles as well that adults are known to rely on. Nisbett and his colleagues (Nisbett, Krantz, Jepson, & Kunda, 1983) have shown that adults are quite willing to infer new information from one member of a category to other category members, but they do so selectively, depending on the property involved. Adults rely on their conceptions of how variable the property is within the domain being questioned to make their inferences. So, their inferences within a category are largely

governed by their theories about the properties involved. Our studies have not tested the limits of children's abilities, but even so it is clear that children must develop more finely tuned distinctions among predicates and learn not to overgeneralize to inappropriate predicates.

Another way in which children are likely to be limited is that they may not be able to look much beyond perceptual features of objects when they form categories on their own. In the present studies, children were told the category labels and then asked to infer information from one member to another. This task is simpler than the converse problem of having to form the category in the first place without knowing beforehand which properties are relevant. When initially forming a category, children are likely to be much more influenced by perceptual appearances. Most standard classification procedures (e.g., Inhelder & Piaget, 1964) require children to divide objects into categories where many bases of classification are possible. Given the complexity of this problem (cf. Gelman & Baillargeon, 1983; Markman & Callanan, 1983), children often find the superficial perceptual appearances of objects to be an easier way of organizing the material.

As for natural kinds, Keil (in press) found young children to be more dependent on perceptual similarity than on deeper biological properties when they are asked to classify anomalous objects. This is again what one would expect, given the limited biological knowledge of young children. Children were asked to classify artifacts and natural kind objects, given conflicting information. For example, one object looked exactly like a skunk but its biological functions (e.g., heart, bones) and lineage (parents) were supposed to be that of a raccoon. The youngest children believed for both natural kinds and artifacts that appearance determined category membership. For example, they would say that the animal that looked like a skunk was in fact a skunk, even though they had been told it had a raccoon heart, gave birth to raccoon babies, and so forth. Not until about second grade were children willing to say that internal structure was an important criterion for categorizing natural kind objects. But younger children probably have no way of knowing whether internal structure or external structure is more important. Another way of stating the difference between Keil's task and ours, is that children were asked to make different sorts of comparisons in the two sets of studies. On Keil's task, children had to compare two different kinds of attributes — perceptual appearance versus biological properties — and determine which is more important for defining the category. On our task, children had to compare the relative importance of category membership versus perceptual similarity. Recent work by Gelman, Collman, and Maccoby (1986) demonstrates that these two tasks are reliably different for young children.

Finally, as Carey (1985) has argued, children must learn how natural kind

categories are related to one another in a system of theory-based knowledge. Carey has found that children initially organize biological knowledge around humans as a prototype. Inferences about the biological properties of other species are based both on what children believe about people and on how similar the species is to humans. For example, children were taught that people (or some other animal) had an omentum and were asked whether various animals and artifacts had an omentum. Children drew inferences about biological categories primarily when the property was known to be true of people. Children were more likely to infer that a biological property of *people* would generalize to bugs than they were to infer that a property of *bees* would generalize to bugs, despite the far greater similarity of bugs to bees. This marked dependence on humans as the prototype changes with age: adults generalize from one species to another based on how similar the species are to each other (Rips, 1975).

Beyond age 4 then, children have much to learn about the relationship between categories and induction. Yet at an age when children are known to find perceptual appearances compelling and when their underlying theories are impoverished at best, they nevertheless expect rich similarities among natural kind objects with the same name. Perhaps 4-year-olds have gathered enough information for them to have reached this conclusion about the structure of categories on an empirical basis. It is also possible that children are initially biased to interpret category terms this way, independent of experience. Other expectations about the structure of natural language categories appear quite early. When children as young as 18–24 months hear an object labeled with a common noun, they assume the term refers to the object as a whole rather than to one of its properties (Macnamara, 1982). By 3 or 4 years of age and possibly earlier, children expect a noun to refer to objects that are taxonomically related (e.g., a dog and a cat) even though in the absence of a label they are likely to group objects on the basis of thematic relations (e.g., a dog and a bone) (Markman & Hutchinson, 1984). The assumption that categories promote rich inductive inferences could be another early bias, one that helps children acquire category terms rapidly, organize knowledge efficiently, and induce information to novel exemplars of familiar categories. By expecting unforeseen nonperceptual properties to be common to members of a kind, children could go beyond the original basis for grouping objects into a category and discover more about the category members than they knew before. Children might start out assuming that categories will have the structure of natural kinds. With development children would then refine these expectations, limiting them to properties, domains, and category types that are appropriate.

References

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: Bradford Books.

Dickinson, D.K. (1982). The development of children's understanding of materials: A study of theory construction and conceptual development. Unpublished doctoral dissertation, Harvard University.

Flavell, J.H. (1963). *The developmental psychology of Jean Piaget.* Princeton, NJ: D. Van Nostrand.

Flavell, J.H. (1985). *Cognitive development,* 2nd ed. Englewood Cliffs, NJ: Prentice Hall.

Flavell, J.H., Flavell, E.R., & Green, F.L. (1983). Development of the appearance-reality distinction. *Cognitive Psychology, 15,* 95–120.

Gelman, R., & Baillargeon, R. (1983). A review of some Piagetian concepts. In J.H. Flavell & E.M. Markman (Eds.), *Cognitive development,* Vol. 3 of P.H. Mussen (General Ed.), *Handbook of child psychology.* New York: Wiley.

Gelman, S.A. (1984). Children's inductive inferences from natural kind and artifact categories. Unpublished doctoral dissertation, Stanford University.

Gelman, S.A., Collman, P., & Maccoby, E.E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development, 57,* 396–404.

Goodman, N. (1955). *Fact, fiction, and forecast.* Cambridge, MA: Harvard University Press.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child.* New York: W.W. Norton.

Keil, F.C. (1979). *Semantic and conceptual development: An ontological perspective.* Cambridge, MA: Harvard University Press.

Keil, F.C. (in press). The acquisition of natural kind and artifact terms. In W. Demopoulas & A. Marras (Eds.), *Language learning and concept acquisition.* Norwood, NJ: Ablex.

Kripke, S. (1971). Identity and necessity. In M.K. Munitz (Ed.), *Identity and individuation.* New York: New York University Press.

Kripke, S. (1972). Naming and necessity. In D. Davidson & G. Harman (Eds.), *Semantics of natural language.* Dordrecht: D. Reidel.

Macnamara, J. (1982). *Names for things: A study of human learning.* Cambridge, MA: MIT Press.

Markman, E.M., & Callanan, M.A. (1983). An analysis of hierarchical classification. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 2). Hillsdale, NJ: Lawrence Erlbaum Associates.

Markman, E.M., & Hutchinson, J.E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic vs. thematic relations. *Cognitive Psychology, 16,* 1–27.

Mill, J.S. (1843). *A system of logic, ratiocinative and inductive.* London: Longmans.

Nisbett, R.E., Krantz, D.H., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90,* 339–363.

Putnam, H. (1970). Is semantics possible? In H.E. Kiefer & M.K. Munitz (Eds.), *Language, beliefs, and metaphysics.* New York: State University of New York Press.

Rips, L.J. (1975). Induction about natural categories. *Journal of Verbal Learning and Verbal Behavior, 14,* 665–681.

Schwartz, S.P. (1977). Introduction. In S.P. Schwartz (Ed.), *Naming, necessity and natural kinds.* Ithaca, NY: Cornell University Press.

Schwartz, S.P. (1979). Natural kind terms. *Cognition, 7,* 301–315.

Tversky, B. (1985). The development of taxonomic organization of named and pictured categories. *Developmental Psychology, 21,* 1111–1119.

Résumé

Une des fonctions principales des termes d'espèces naturelles (comme, par exemple, *tigre* et *or*) est de permettre de formuler des inférences inductives. On s'attend à ce que les membres de telles catégories partagent des propriétés importantes et imprévues, comme par exemple le type des organes internes et la structure génétique. De plus, il est possible d'effectuer des inférences inductives alors que l'information perceptuelle est fragmentaire, par exemple lorsqu'un objet ne ressemble pas beaucoup à d'autres membres de sa catégorie, ou lorsqu'une propriété est inobservable. Dans ce travail, la question qui nous intéresse est de déterminer l'origine de cette attitude vis-à-vis des espèces naturelles. Les jeunes enfants, qui s'appuient habituellement pour leurs inférences sur les apparences perceptives et n'ont qu'un savoir scientifique rudimentaire, pourraient bien ne pas généraliser de nouvelles informations par induction dans le cas des espèces naturelles. Pour tester cette possibilité, nous avons opposé appartenance catégorielle et similitude perceptuelle dans une tâche d'induction. Par exemple, des enfants devaient décider si un requin respire comme un poisson d'aquarium (tous les deux sont des poissons) ou comme un dauphin (apparence semblable). A partir de 4 ans au plus tard, les enfants peuvent utiliser les catégories pour fonder une inférence inductive même lorsque l'appartenance catégorielle est en conflit avec les apparences. En plus, ces jeunes enfants ont partiellement effectué une séparation entre les propriétés qui autorisent l'induction dans une même catégorie (par exemple la respiration) et les propriétés qui dépendent en fait des apparences perceptuelles (comme le poids). Puisque nous n'avons étudié que le cas des catégories d'espèces naturelles, nous ne savons pas jusqu'à quel point les enfants différencient les espèces naturelles des autres types de catégories. L'hypothèse de départ des enfants pourrait être que toutes les catégories linguistiquement nommées ont la structure des espèces naturelles, et ces attentes pourraient se modifier au cours du développement.