ELSEVIER

Cognition 59 (1996) 307–335

COGNITION

Drinking and driving don't mix: inductive generalization in infancy

Jean M. Mandler[a,b]*, Laraine McDonough[a]

[a]Department of Cognitive Science, University of California, San Diego, La Jolla, CA 92093-0515, USA

[B]Cognitive Development Unit, University College London, 4 Taviton Street, London WC1H 0BT, UK

Received 12 October 1994, final version accepted 29 September 1995

Abstract

The traditional view of inductive generalization in infancy is that it rests on perceptual similarity; infants are said to form perceptually based categories, such as dogs and cats, and then to associate various properties with them. Superordinate-level inductions, such as generalizations about animals as a domain, have been considered to be more abstract and assumed to be a later achievement. Three experiments were conducted to investigate these issues, using 14-month-olds as subjects. We modeled various properties or actions appropriate to animals or to vehicles and then assessed whether infants were willing to generalize their imitations of these actions to different exemplars from the same and different domains. Contrary to the traditional view, we found that infants this age have generalized the properties of drinking and sleeping throughout the animal domain, and the properties of "being keyed" and "giving a ride" throughout the vehicle domain. These generalizations are constrained solely by the boundaries of the domains themselves and are not influenced by the perceptual similarity of exemplars within the domains.

1. Introduction

Our current research is exploring the inductive generalizations that 1-year-old infants make about the classes of animals and vehicles. We ask

* Corresponding author at: Department of Cognitive Science, University of California, San Diego, La Jolla, CA 92093-0515, USA.

SSDI 0010-0277(95)00696-6

whether infants this age have generalized properties such as eating and sleeping across these broad domains. We are interested in what constraints, if any, there are on these generalizations. We assume that 1-year-olds have observed people and perhaps a few other animals eat and sleep. How far have they generalized those observations and what are the limits on them? For example, would generalizations made from observing a dog and a cat eating be limited to those subclasses, or would they be generalized to the entire class of animals, or even to all objects? Similar questions can be asked about knowledge of vehicles. It has been suggested that for older children categories of artifacts engender fewer inductions than do biological categories (e.g., Gelman, 1988). Is this also true for infants? That is, do infants treat biological kinds differently from artifacts?

We know that infants will generalize responses learned to one object (a toy, for example) to another example of the same object that varies only in color or pattern (Baldwin, Markman, & Melartin, 1993). But there is little information in the literature to tell us what generalizations infants have or have not made about more varied classes of objects. Some information can be gleaned from studies of infant imitation. Killen and Uzgiris (1981) showed that 10- and 16-month-old infants are more likely to imitate appropriate actions on objects than counterconventional ones; for example, they are more likely to put a comb to their hair than to a car in response to a model's showing them these two gestures. Bauer and Dow (1994) showed that 16- and 20-month-olds are willing to generalize the props used in event sequences to other members of the same class; for instance, at these ages children will use Big Bird and a toy bed to reproduce an event sequence in which a teddy bear had been put in a toy crib. This study in particular suggests that by 16 months infants have associated certain behaviors such as sleeping with fairly broad classes. However, virtually all the research specifically designed to study how category knowledge limits induction has been conducted with older children. The youngest age at which the extent of inductive generalization has been tested is $2\frac{1}{2}$ years (Gelman & Coley, 1990), and most work on induction has been conducted with children 3 years or older (e.g., Gelman & Markman, 1986, 1987). Not surprisingly, these studies have all involved verbal information. For example, children are told some fact about an object and are then asked to decide the categorical range over which the taught property is valid (Carey, 1985; Gelman & Reilley, 1988).

Since the present experiments study preverbal infants we cannot ask our subjects about their categorical knowledge, so we make use of nonverbal techniques instead. Using little models of objects we demonstrate a property or action characteristic of a given class and then test whether the infants will generalize the imitation of that action to other exemplars from the same class and/or those of a different class. This technique provides us with information about infants' conceptual categories in the sense that it tells us where the boundaries of given classes lie and the sorts of generalizations infants are willing to make across these classes. It does not, of course, tell us about their knowledge in the explicit way that verbal questioning does.

1.1. *Do infants generalize on the basis of similarity or kind?*

In order to understand how infants begin to form inductive inferences, it is important to distinguish between their capacities to form perceptual categories of objects that look alike and to form conceptual categories of objects that are the same kind of thing. Our previous research has shown that before 1 year of age infants have already formed global conceptual categories, or concepts, of animals and vehicles (Mandler & McDonough, 1993). We say they have formed concepts of animals and vehicles because in various tasks they treat all animals as if they were the same and all vehicles as if they were the same in spite of the fact that most members of these large superordinate classes do not look alike. We do not yet know, however, exactly what these concepts consist of for infants or whether they serve the same functions that concepts do for older children and adults. One of the main functions of conceptual categories or concepts is to allow inductive generalizations across the exemplars of a class, a function that allows a wide range of knowledge to be acquired from a limited set of observations. However, it might be the case that the early conceptual categories that infants form are too vague or unstable or have boundaries too hazy to provide an adequate foundation for generalization. In this respect they might be closer to what Piaget (1951) called "preconcepts". Alternatively, it might be that these first conceptual classes are stable enough to support generalizations but that infants do not yet make conceptual generalizations, at least not about such large and varied classes as animals or vehicles.

The traditional empiricist doctrine, espoused in modern times by philosophers such as Quine (1977), is that the first inferences are *not* based on a conceptual system of kinds, but instead are generalizations based on innate responsivity to physical similarity. Physical similarity itself is described in terms of sensory qualities, such as color, shape, and texture, that can be directly perceived without being mediated by a conceptual system. Quine (1977) called this early responsivity an immediate, subjective, animal sense of similarity. According to this doctrine, which Keil (1991) has dubbed the doctrine of "original sim", before children develop abstract concepts or theories about the world, they are only influenced by the laws of similarity; they make associations and generalizations on the basis of what things look like. The more two things resemble each other the more likely it is that an inductive inference from the properties of one to the other will be made. For example, an infant may see a cat eat, and from this observation (or perhaps several such observations) induce the generalization that all cats eat.

In this view, the first categories to be acquired are the so-called basic-level categories (Mervis & Rosch, 1981), formed on the basis of the physical similarity of one object to another: cats have very similar shapes (and presumably texture and movement), all of which are very different from the shapes (and texture and movement) of fish or tables. Once these simple perceptual categories or perceptual schemas are formed, infants can begin to associate them with various sounds, names, and typical actions. It appears that this is assumed to be the process by which a perceptual category or schema becomes a basic-level concept: various properties become associated with a perceptual core (Eimas, 1994). Therefore, in this view, induction of properties first takes place at the basic level, where perceptual similarity among exemplars is high. With more experience, the process of induction gradually changes, eventually resulting in generalization at a more abstract level. Cats eat, fist eat, birds eat; therefore, even though they have different shapes, texture, and movement, perhaps all animals eat. This kind of induction is said to be a more difficult accomplishment because of the lack of similarity among the members of superordinate classes (Mervis & Rosch, 1981).

Consistent with this view is the fact that infants do form perceptual categories or perceptual schemas of real-world objects at a very early age. Using realistic pictures of animals, Quinn, Eimas, and Rosenkrantz (1993) have shown that within a few trials 3-month-olds form a perceptual category of dogs that excludes cats, and Eimas and Quinn (1994) have shown that the same-aged infants also form a category of horses that excludes zebras and a category of cats that excludes lions. At this age, these categories appear to be purely perceptual achievements based on the physical characteristic of shape, without any conceptual implications. That is, infants are able to categorize dog-like patterns or cat-like patterns without having any ideas as to what kinds of things these are and without having associated any properties with them, such as that one of these objects barks and the other meows. These are the categories that people apparently have in mind when they say that basic-level categories are the first to be formed. If early inductive inferences are based on physical similarity, then these perceptual categories should form the foundation on which inductive inferences rest.

There are a number of difficulties with this approach to concept formation and induction of concept properties. First, it does not specify the kind of perceptual similarity that matters in forming the earliest concepts. Traditionally, arguments from similarity have called upon likeness in shape to explain the formation of basic-level classes (e.g., Mervis & Rosch, 1981), although occasionally other dimensions such as texture or movement have been suggested as playing a role. Domain-specific features such as eyes or wheels are less often called upon, because these do not distinguish most basic-level classes within a domain. Everyone does agree that basic-level exemplars have a higher degree of similarity (however defined) than do superordinate exemplars, because the latter necessarily have less overlap in features, and often little overlap at all (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Nevertheless, because this position has not been required to specify precisely what constitutes perceptual similarity, it is always possible to claim that some feature, such as eyes or wheels, is so salient that it can override shape, texture, or movement differences and make all animals (or all vehicles) perceptually "similar". It should be noted, however, that defining similarity in terms of special features requires some kind of (presumably innate) bias; perceptual similarity itself cannot explain why infants pay attention to one kind of feature rather than another.

Second, the argument for the primacy of perceptual similarity does not specify how anything not purely perceptual is represented in infancy, so that it can be associated with a given perceptual category, such as dog or cat. A feature such as barking is straightforwardly perceptual, so one can imagine how it might be associated with the perceptual category of dogs. But how does the infant represent less obviously perceptual characteristics such as "eats," "interacts with me," or "starts up on its own"? These characteristics can be observed, of course, but they are not purely perceptual in the sense of consisting solely of distinctive shape, texture, motion, and so forth; they involve meaning. For example, putting a spoon to one's mouth is physically more similar to putting a spoon an inch to the right of one's mouth than it is to putting a bottle to one's mouth. But the meaning of eating makes us judge that the spoon and the bottle actions are alike, not hitting or missing the mouth with the spoon. It is this kind of meaning that is difficult to represent in purely perceptual or sensory terms.

Third, an approach based on physical similarity cannot explain how more conceptually based inductive generalizations begin to take place. What enables the infant to go beyond generalizations about cats to generalizations about animals? One of the reasons that superordinate concepts are said to be late acquisitions is just because they require sensitivity to nonobvious, less perceptually based (theoretical) properties. However, to say that responding to nonobvious properties is difficult does not say how the difficulty is overcome. As both Quinn and Eimas (1986) and Keil (1991) have pointed out, no one espousing the traditional view has ever shown satisfactorily how generalization on the basis of perceptual similarity gets replaced by theory-based generalization. Yet this transformation must have taken place by age $2\frac{1}{2}$, since it has been shown that by this age conceptual class membership overrides perceptual similarity in determining induction (Gelman & Coley, 1990). Thus, the traditional doctrine not only has difficulty explaining how abstract or superordinate categorization develops, it has equal difficulty in explaining how conceptually based inference supersedes similarity-based inference.

An alternative hypothesis is that, even in infancy, making inductive inferences is based not on the physical appearance of things but instead on the notion of kinds. Little work has been carried out to study how kinds are conceived in infancy. However, we assume that even for infants the notion of a kind is based on meaning, rather than on physical appearance per se. The research in our laboratory suggests that forming conceptual categories or kinds proceeds in a different fashion from forming uninterpreted perceptual categories. The work of Eimas, Quinn, and colleagues indicates that early perceptual categorization of animals rests on sensitivity to shape and facial patterning (Quinn & Eimas, 1994), and perhaps to individual perceptual features when they are contrastive (for example, differentiating horses from zebras on the basis of their stripes). Forming concepts of things, on the other hand, does not seem to depend upon individual perceptual features or even upon the overall perceptual similarity of the exemplars involved.

As early as 7 months, infants have begun to categorize animals as different from vehicles, in spite of the great perceptual variability in the exemplars within each of these domains (Mandler & McDonough, 1993). At the same time, infants make few conceptual distinctions *within* the animal domain. For example, we give 7- to 11-month-olds a series of little models of animals to examine one after the other (Mandler & McDonough, 1993). The animals vary considerably in their appearance and infants treat each one with equal interest as measured by examination time.[1] If we then give them a model of a vehicle, their examination time increases markedly. On the other hand, when they are given a series of models of dogs to examine, they examine each one with roughly the same interest, but do not show any differential response when they are then given a model of a fish or a rabbit.[2]

This lack of responsivity to the difference between dogs and fish and rabbits is not due to the task being a perceptually insensitive one. The same infants differentiate birds and airplanes even though all of them are very similar in shape, including outstretched wings (Mandler & McDonough, 1993). Instead, the data suggest that the infants are treating dogs and fish and rabbits as if they were all the same kind of thing, even though they perceive the differences in their appearance. Seeing the difference between dogs and rabbits but treating them as the same kind of thing is similar to an older child seeing the difference between subordinate classes, such as collies and poodles, yet treating them both as the same kind of thing – namely, dogs. This lack of behavioral responsivity to perceived differences suggests that at a conceptual level the main distinction 7-month-olds are making is that between animal and nonanimal kinds. This is not a particularly surprising finding; 7-month-olds in our culture typically have had relatively little experience with different kinds of animals. At the same time, if they have observed any animals at all they will have experienced the characteris-

[1] The sets of animals and vehicles we have used are as varied as possible; the animals include fish, birds, and mammals and the vehicles include motorcycles, train engines, and helicopters. Some of the animals show no visible eyes or mouths; some of the planes have facial features (as in Flying Tigers). The models are all either plastic or metal, so between-class textural differences are minimized.

[2] The only differentiation 7-month-olds have shown among subclasses of animals on the object-examination task is between dogs and birds.

tics that differentiate animals as a global class from nonanimals, such as that they move on their own and interact with other objects from a distance (Mandler, 1992). These characteristics, however, are abstract and not tied to the kind of specific perceptual details that most people have in mind when they call upon perceptual similarity. Another way of expressing this distinction is to say that infants readily conceptualize animals as a class of things different from vehicles but find it more difficult to conceptualize how one animal differs from another.

Progress in conceptually differentiating animal kinds occurs over the next 10 months. For example, current work in our laboratory indicates that 7-month-olds do not treat dogs and cats as different kinds in the object-examination task but by 11 months they do. The latter result is 8 months after the age at which they can *perceptually* categorize these classes. With older infants we use a task that is closer to the classifying tasks used with older children; we put in front of infants a number of exemplars from two different classes and observe the order in which they touch them. On this task, by 18 months infants make a tripartite division of the animal world into land-, air-, and water-animals (Mandler, Bauer, & McDonough, 1991), although they still do not separate dogs from horses or rabbits. The literature on language acquisition indicates, however, that by 16 to 18 months they are beginning to attach appropriate labels to members of animal subcategories, and to learn some properties of these smaller classes, such as characteristic sounds. Although they often overextend these labels to other animals (e.g., Clark, 1983), conceptual differentiation of the animal domain into so-called basic-level concepts does seem to increase around this age.

This brief summary indicates that the global differentiation of animals from nonanimals as distinct kinds of things begins to develop early but that individuation of the animal domain into distinct kinds is a slower process. Although there is a lesser amount of data on vehicles, we have found a similar pattern of differentiation, with the exception that in one of our studies infants showed responsivity to subclasses of vehicles as young as 7 months (Mandler & McDonough, 1993).[3] These data also indicate that there are differences in the course of perceptual and conceptual development. The very slow and gradual differentiation of large, globally specified classes into more and more detailed subclasses does not fit easily with data such as those of Eimas and Quinn (1994) showing the rapid acquisition of perceptually detailed subclasses such as horses and zebras at 3 months of age. It does not seem possible that both developments can be explained by a single process of responsivity to perceptual similarity.

It is important to stress that we are not saying that perceptual information

[3] This has not been true in all our experiments. In Mandler et al. (1991) there were no differences in the course of differentiation within the animal and vehicle domains. For example, at 18 months, infants were still not separating cars from trucks or motorcycles.

is unimportant to concept formation. On the contrary, we believe that it is *analysis* of perceptual information that provides the initial foundation for conceptualizing the kinds of things there are in the world, that is, for giving perceptual categories meaning. In particular, we have suggested that perceiving the way that objects move is crucial for understanding what they do and therefore for conceptualizing what they are (Mandler, 1992). Although the actual movements that animals make vary greatly from species to species, at a more abstract level, these movements have several characteristics in common; for example, they all begin movement on their own and interact contingently with other objects from a distance. Thus, there is responsivity to sameness or alikeness both in forming a perceptual category and in forming a conceptual class, but in the one case it is alikeness in appearance and in the other it is alikeness in meaning. In addition, of course, infants must use perceptual features whenever they identify an object as a member of a given conceptual class. Our claim is only that what holds a class such as animals together so that membership can be assigned to it is neither overall perceptual similarity nor a common perceptual feature, but instead a notion of common kind.

If it is accepted that infants can have both detailed perceptual categories such as dog or horse and at least global conceptual categories of kinds, it becomes reasonable to ask which they use for their earliest inductions. Do infants rely on their perceptual categories for inductive inference or on their conceptual categories? Since our data indicate that infants are not restricted to responding to animal similarity, it may be that they can use their beginning concepts for purposes of generalization. If so, a qualitative shift in induction from animal similarity to class membership need not be required.

1.2. Does similarity play a role in the earliest inductions?

Even though class membership is considered to be the most important factor in constraining induction (e.g., Gelman & Markman, 1986; Quine, 1969), research on both children and adults has generally supported the notion that the similarity between trained and test exemplars plays at least some role (Carey, 1985; Gelman, 1988; Rips, 1975). Thus, within domains, the more similar a test exemplar is to the trained exemplar, the more likely people are to generalize from one to the other. For example, Carey (1985) showed that generalization of animal properties such as sleeping or having a heart falls off as the test exemplars become less and less like humans. All 4-years-olds attributed attributes such as these to people, but fewer attributed them to an unfamiliar mammal (aardvark), and fewer still to a shark, insect, or worm. A similar, although less pronounced, drop-off was found for 10-year-olds and adults. Similarly, Rips (1975) showed that within the category of birds, adults were more likely to generalize from eagles to hawks than to ducks, and from ducks to geese than to eagles. Gelman (1988) and Gelman and O'Reilley (1988) showed that children and adults were most apt to make inductive generalizations to other members of the same subordinate class, next most likely to members of the same basic-level class, less likely to generalize to other members of the same superordinate class, and made few if any inductive generalizations to members of an entirely different domain.

This type of similarity is sometimes called category homogeneity (Gelman, 1988); the more homogeneous the exemplars of a given class are, the more likely one is to generalize from one member to another of that class. In general, of course, the smaller the subclass in a taxonomic hierarchy, the more homogeneous its members. Thus, exemplars of a particular type of dog, such as poodles, are more similar in appearance than are exemplars of the larger class of dogs. And dogs are more similar in appearance to each other than are the exemplars of the larger classes of mammals or animals. Thus, there is a natural (although not perfect) confound between height of a class in a taxonomic hierarchy and dissimilarity. This makes it difficult to know in the various studies described above whether the drop-off in generalization from one class to another occurs because of decreasing similarity of the exemplars or from moving farther away in a hierarchically structured set of classes.

Some work by Gelman (1988) suggests that similarity unconnected to class membership may be less important than the breadth of the class in controlling induction. As just discussed, category homogeneity (similarity) as defined by level in a taxonomic hierarchy had a major effect on likelihood of making inductive inferences. However, Gelman also examined the ratings of category homogeneity within classes, all of which were at a given level (the basic level was used for this analysis). Some basic-level classes have exemplars that look more alike than those of other basic-level classes. Although Gelman did not report the ratings of homogeneity for the various basic-level classes she used, one can assume that a class such as carrots had members that looked more alike than a class such as flowers. In any case, no relationship was found between similarity of exemplars and inferences in the natural kinds that she studied (although a modest relationship was found for artifact categories). This finding suggests that it may be the breadth of the conceptual class that is controlling induction even for preschool children more than similarity of the exemplars within a given class. Smaller classes may be more conducive to induction, not because their members look alike, but because there is less conceptual variation associated with them.

As far as infants are concerned, we have seen that they do not use overall perceptual similarity in their early conceptual categories. Therefore, it is plausible that they might not use this factor when making inductions either. If properties such as self-motion or interaction from a distance form the core notion of animal, and if that is all that infants have conceptualized about animals at the time they begin making inductions across this domain, then the overall physical similarity of various exemplars may not play much of a role in constraining the types of induction that infants make. As just suggested, the value of similarity of appearance is that it suggests that two things are the same kind. The more subcategories one has formed in a domain, the more useful perceptual similarity is in determining the smallest class, which in turn is useful for purposes of inductive certainty. But if one has made few if any subdivisions in a large and variable domain such as animals, then perceptual similarity of the exemplars to each other may be of relatively little use.

Based on these considerations as well as on the prior work with preschool children, we hypothesized that even the earliest inductive generalizations would be determined by conceptual relatedness more than by similarity of the old and new exemplars, and that the scope of the generalizations infants make would be determined by the smallest classes they have conceptualized. Since our previous work indicated that the concepts of animals and vehicles are still very global for 14-month-olds (the age of the subjects in the present experiments), we predicted that generalization from old to new instances would be general across an entire domain.[4] Therefore, there should be little if any influence of basic-level classes (or any other subclasses) on likelihood of generalization, because these subdivisions of the animal and vehicle domains have not yet been conceptualized as different. Thus, we expected to find equal amounts of generalization no matter which instances of the domain were used.

These hypotheses were tested in three experiments. In the first experiment we modeled for 14-month-olds an action appropriate either to the animal or vehicle domain. For the generalization test, we gave them a different exemplar from the target domain and an exemplar from the other domain, along with the prop used to carry out the action (e.g., a cup or a key), and observed which exemplar the infants used for their imitations. The new exemplar from the target domain was either similar or dissimilar in appearance to the modeled exemplar. In the second experiment we used exotic exemplars from each of the domains for the generalization tests to ensure that the subjects would have never seen them before. The third experiment tested the limits on the infants' generalizations by modeling both appropriate and inappropriate actions on animals and vehicles and then giving the infants the same objects for their imitation to assess whether they would be willing to imitate the inappropriate actions that they had seen modeled. This experiment also acted as a control on the extent to which the infants were generalizing play behavior to "toys" rather than generalizing the behavior of the objects the toys represent. If the infants were only generalizing the experimenter's play behavior without any implications for

[4] Although much work remains to be done on domain boundaries, we currently have enough data to indicate that by this age, infants are not merely making a distinction between animate and inanimate things, but at the least are making distinctions among the domains of animals, vehicles, plants, and household items.

the world at large, then one would expect no differential tendency to imitate appropriate versus inappropriate behavior.

EXPERIMENT 1

2. Method

2.1. Subjects

Thirty-two infants were seen when they were 14 months of age (mean = 14 months, 13 days; range = 14 months, 0 days to 14 months, 27 days). Subjects were randomly assigned to one of two groups: a similar test group and a dissimilar test group. In each group, the same number of females and males were tested. Subjects were recruited from an existing pool of volunteer parents who had responded to advertisements in local newspapers. A small toy was given to the infants for their participation. All subjects who were tested on both days were included in the analyses.

2.2. Properties tested

Each subject was tested on four properties: two appropriate to animals and two appropriate to vehicles. The animal properties were drinking from a cup and sleeping. These were modeled by putting a little cup to the face of a small replica of an animal and by putting a small replica of an animal in a little bed. The vehicle properties were giving a child a ride and using a key on a vehicle. The first of these was modeled by placing a doll, who was in a seated position, on top of a replica of a vehicle and scooting it. The second was modeled by touching a key to the side of a vehicle. Table 1 lists the exemplars, the properties tested, the props used for modeling the actions, and the vocalizations that accompanied the modeling.

All the exemplars were small replicas of animals and vehicles, shown in Fig. 1. The animal properties were modeled with a prototypical dog (German shepherd) and the vehicle properties were modeled with a prototypical car (a sedan). These two exemplars were used only for modeling the properties and were not given to the subjects. Instead, subjects were tested using other animals and vehicles (the test exemplars). For the subjects in the similar test group, the test animals were similar to the dog (a rabbit and a cat) and the test vehicles were similar to the car (a truck and a bus). For the subjects in the dissimilar test group, the animals and vehicles were dissimilar to the dog (a fish and a bird) and car (a motorcycle and an airplane). For each property, subjects were tested by presenting them with a target exemplar from the modeled domain and a distractor exemplar from the other domain. The similar exemplars served as distrac-

Table 1
Exemplars, actions, props and vocalizations used in Experiment 1

		Similar group ($N = 16$)		Dissimilar group ($N = 16$)	
Exemplar for modeling action	Actions, props and vocalizations	Similar test	Distractor	Dissimilar test	Distractor
Dog	Drinking from a cup "sip, sip, umm, good"	Rabbit or cat	Motorcycle or airplane	Bird or fish	Truck or bus
	putting in a bed and patting "night night"	Cat or rabbit	Airplane or motorcycle	Fish or bird	Bus or truck
Car	Starting with a key "vroom, vroom"	Truck or bus	Bird or fish	Motorcycle or airplane	Rabbit or cat
	Giving a seated child a ride "go for a ride, wheee"	Bus or truck	Fish or bird	Airplane or motorcycle	Cat or rabbit

tors for the subjects in the dissimilar test group and the dissimilar exemplars served as distractors for the subjects in the similar test group. (Because of this control, the two most similar cross-domain exemplars, namely, the bird and airplane, were never directly paired together as target and distractor. This pairing was used in Experiment 3, however.) Order and assignment of targets and distractors were counterbalanced within each group so that each exemplar was used to test each property equally often.

2.3. Procedure

Subjects were invited into a laboratory set up as a playroom. After a brief warm-up period they were seated in the parent's lap across the table from the experimenter. Parents were asked not to assist their infants in any way throughout the sessions. Two warm-up tasks, designed to accustom the subjects to the imitation procedure, were administered prior to the generalization task. Both were exercises in hammering an object. In the first task, one pipe was fitted into another and a plastic hammer was used to pound them tightly together with the accompanying vocalization "bam, bam, bam, all fixed!" In the second warm-up, the experimenter but two large Lego pieces together and hammered them with the same vocalization. Subjects were encouraged to imitate both actions and were praised for doing so. All but one subject imitated both the warm-up actions.

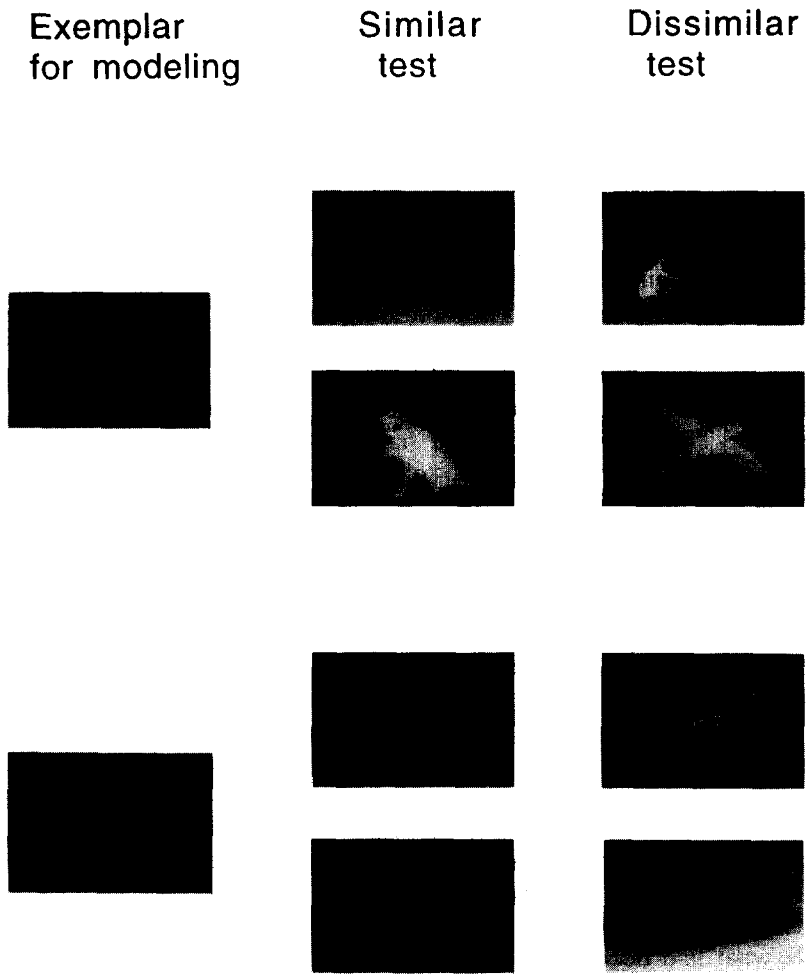

Fig. 1. The exemplars used for modeling and the generalization tests in Experiment 1.

Because the procedure required handling many different objects, subjects were tested during two sessions on two separate days. In the first session, one of the animal and one of the vehicle properties were tested. During the following session the remaining animal and vehicle properties were tested. Order in which the properties were tested was counter-balanced across subjects. The same warm-up tasks were administered at the beginning of each session.

Subjects' demonstrations of the properties were evaluated twice during each session, once before the properties were modeled (baseline) and again after they were modeled (generalization).

Baseline: before the properties were modeled, subjects were given both test exemplars (target and distractor) and the prop used to demonstrate the target action. The amount of time allotted for baseline was determined by the subjects. If a subject ignored any of the three objects, the experimenter would point it out saying "Look, did you see this one?" After the subjects were finished exploring each of the objects, the experimenter took them away. The prop was left on the table and the test exemplars were placed out of sight. The experimenter then brought out the modeling exemplar (e.g., dog) and demonstrated the target action with the prop (e.g., a bed) three times with an appropriate vocalization (e.g., "night, night"). The experimenter then removed the dog from the table so that it could no longer be seen.

Generalization test: the test exemplars were brought back and simultaneously placed to the right and left of the subject. The target exemplar was placed equally often to the subject's right and left throughout the two sessions. The experimenter then handed the prop to the subject while repeating the vocalization (e.g., "night, night"). This procedure continued throughout both sessions until each subject had been tested on both the animal and both the vehicle properties.

2.4. Scoring

Each session was videotaped in this and the following two experiments. Subjects were coded for performance (or nonperformance) of the properties and the exemplar they used for doing so. If a subject used both test exemplars to demonstrate the properties, the coder noted which was chosen first. The agreement between two coders in the three experiments ranged from 95% to 100%. The few disagreements were resolved by a third coder. Coders were blind to the hypotheses of the experiment, but not to the acts that were modeled (they were obvious given the props that were being used and the infants' own actions during the test). Because it was not possible to hide what had been modeled from the coders, a fourth coder was given a misleading hypothesis and asked to score a sample of the tapes. She was told that engaging in counterconventional acts is the most sophisticated level of imitation and that she should be particularly alert for any such occurrences. However, these instructions to look for actions being performed on the distractors did not produce any differences in scoring; her agreement with the main coder was 96%.

3. Results

For the first analysis, subjects' first choices for demonstrating the target properties were analyzed in a mixed-design analysis of variance. The between-factors were similarity (similar, dissimilar) and gender. The within-factors were domain (animal, vehicle), exemplar (target, distractor) and assessment (baseline, generalization). The dependent variable was the percentage of properties demonstrated by each subject.

The results showed that infants demonstrated significantly more of the properties using the target exemplars ($M = 49\%$) than the distractor exemplars ($M = 11\%$), $F(1, 28) = 62.55$, $p < .001$. In addition, more of the properties were demonstrated during generalization ($M = 39\%$) than during baseline ($M = 20\%$; $F(1, 28) = 60.49$, $p < .001$). However, as depicted in Fig. 2, these two effects are qualified by an interaction between them, $F(1, 28) = 17.52$, $p < .001$.

In Fig. 2 it can be seen that subjects demonstrated the properties with the target exemplars more often than with the distractors both at baseline, $F(1, 28) = 11.45$, $p < .01$, (target $M = 30\%$, distractor $M = 10\%$) and at generalization, $F(1, 28) = 62.28$, $p < .001$ (target $M = 67\%$, distractor $M = 11\%$). At the same time, performance with the target exemplars increased significantly from baseline to generalization, $F(1, 28) = 42.80$, $p < .001$, but did not do so for the distractors, $F(1, 28) = 0.02$. A main effect was also found for domain, $F(1, 28) = 29.70$, $p < .001$, showing that more properties were demonstrated in the animal domain ($M = 36\%$) than in the vehicle domain ($M = 24\%$). There was no significant effect for similarity, $F(1, 28) = 1.07$, although properties were demonstrated somewhat more often on the similar objects ($M = 55\%$) than on the dissimilar objects ($M = 42\%$). This difference in the means stemmed mainly from the relatively small number of times subjects used the airplane to give a ride (discussed below). There were also no significant effects of gender.

Next, we examined the data by combining the first and second choices. The pattern of results was highly similar to those found in the previous

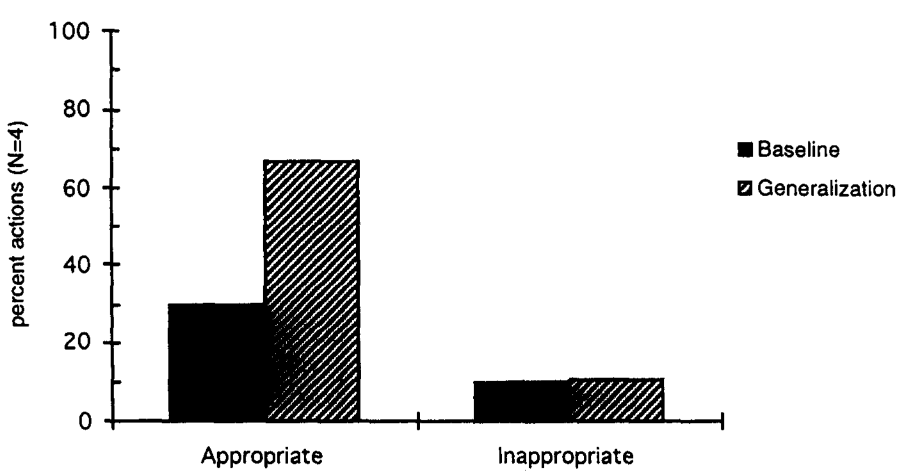

Fig. 2. Experiment 1. Percentage of first actions performed to target and distractor objects.

analysis when only first choices were used. Performance with the appropriate but not the inappropriate exemplars increased significantly from baseline to generalization. We then made a comparison between subjects' first and second choices. On the few occasions when subjects first demonstrated a property with an inappropriate exemplar (a total of 12 occurrences at baseline, 15 at generalization, each of 128 opportunities), they frequently went on to demonstrate the property with the appropriate exemplar as well (33% of the time). However, when their first selections were made to an appropriate exemplar (39 occurrences at baseline, 85 at generalization), they made second choices less often (13.5% of the time). This difference in conditional probabilities approached significance, $t(31) = 1.57$, $p < .07$, one-tailed test. Thus, subjects relatively rarely generalized a property to the inappropriate exemplar after having generalized it appropriately on their first demonstration.

The data were then examined for differences among the four properties tested. Two 2×2 chi-square tests were conducted: one on the two animal properties, the other on the two vehicle properties. Group and property tested were the factors used. The results showed that fewer subjects in the dissimilar test group generalized the riding action ($N = 5$) than the keying action ($N = 11$) relative to the subjects in the similar test group, who generalized both properties equally often (riding $N = 12$, key $N = 12$), chi-square $(df = 1) = 3.81$, $p < .06$. This result was due to the subjects in the dissimilar group not using the airplane to give a ride ($N = 1$), compared to the motorcycle ($N = 4$), truck ($N = 6$), and bus ($N = 6$). No significant differences were found in the generalizations of the animal properties between the similar and dissimilar groups, chi-square $(df = 1) = 0.65$. Further chi-square tests showed no significant differences among the individual exemplars. That is, with the exception of the low generalization of riding the airplane, subjects used all the exemplars equally often for demonstrating the other properties tested.

Finally, a number of instances of spontaneous performance of other domain-appropriate behaviors were recorded both during baseline and generalization. For example, on 10 occasions subjects gave an animal a ride on a vehicle in addition to (or instead of) the test person.

4. Discussion

When 14-month-olds were shown appropriate actions performed on a model of an animal or a vehicle, they were willing to generalize these actions to any other animal or vehicle that was provided. No effect of similarity between the exemplars used for modeling and those used for generalization was found. After seeing a dog drinking from a cup or being put to bed, the infants were as willing to make a bird or a fish do the same as they were to make another land-animal do so. Similarly, after seeing a car started with a key or giving a person a ride, the infants were as willing to start an airplane or a motorcycle as another four-wheeled vehicle. The only exception was that the infants gave the person a ride on an airplane slightly less often than on a truck. The most likely reason for this exception (which was not found in Experiment 3) was some difficulty in making the particular doll that was used sit on top of the airplane. Thus, with the possible exception of giving a ride, the responding was general across the entire domain being tested and not confined to those exemplars that looked most similar to the modeled exemplar. At the same time, the generalization was constrained by the relevant domain boundaries. Infants rarely imitated the modeled action on an exemplar from the wrong domain.

These results were not confined to imitation. Although only 20% of the to-be-modeled actions were demonstrated during baseline, the majority of these were performed on objects from the appropriate domain (30% vs. 10%). This finding is particularly interesting given that the situation was playful and the infants appeared to treat it as such. There was nothing in the situation itself that constrained their actions to be appropriate, yet they were. In addition, both during baseline and generalization a number of the subjects spontaneously demonstrated the properties that we tested.

The relatively high baseline for actions on the appropriate objects (30%) indicates that many (perhaps all) of the infants had already learned and generalized the properties tested. That is, there is no evidence that the various properties were being taught during the experiment. When the generalizations were first made is an interesting question that is important to answer in order to compare these data directly with those on older children (e.g., Gelman, 1988), but it cannot be answered from the present data. The only conclusion that can be made is that at whatever age properties such as drinking and sleeping, or keying and providing rides are learned, they become associated with the global domains to which the observed exemplars belong.

At present we have no information about the course of this learning. Fourteen-month-olds might have observed a dog being given water or fed, or have observed a dog sleeping in a bed. They were even more likely to have observed the particular actions we modeled in the vehicle domain, namely, using a key on a car or a car giving a person a ride. We do not know, of course, what animals or vehicles they might have seen engaging in these actions. However, they were unlikely to have observed the dissimilar test items engaging in these behaviors. That is, it seems unlikely that infants would have observed a bird being given a drink from a cup or going to bed, or an airplane being opened with a key. Nevertheless, in this age of imaginative television programming, such sights are possible, and some parents may have begun to engage in pretend play with toys with their infants by 14 months. Alternatively, the infants might have observed older siblings engaging in such behavior with birds or airplanes. In order to be sure that the infants were indeed making generalizations and not duplicating previously observed behavior on specific objects, we replicated Experiment 1 using as atypical and unusual exemplars from the animal and vehicle domains as we could find.

EXPERIMENT 2

5. Method

5.1. Subjects

Sixteen subjects, 8 males and 8 females, participated in the experiment. Their mean age was 14 months, 11 days (range = 14 months, 1 day to 14 months, 29 days). Subjects were obtained from the same source as in Experiment 1 and were given a small gift for their participation. One additional subject was tested but not reported in the data due to experimenter error.

5.2. Procedure

The same properties tested in Experiment 1 were tested again, this time using atypical animal and vehicle exemplars. As in Experiment 1, baseline was assessed by giving the subjects the opportunity to manipulate the test exemplars and props. The exemplars used to test the animal properties were an armadillo and an anteater (targets); these were contrasted with a tractor and a crane (distractors). The exemplars used to test the vehicle properties were a forklift and a shoveler (targets) which were contrasted with a musk ox and a dinosaur (distractors). Across subjects, each target was paired equally often with each distractor. As in Experiment 1, once baseline was assessed and the exemplars were put away, an animal property was modeled with the dog or a vehicle property was modeled with the car. After the dog or car was put away, generalization was assessed by bringing out the test target and distractor exemplars, placing them to each side of the subject and handing the subject the prop with the accompanying vocalization (e.g., "night night").

As in Experiment 1, subjects were tested in two sessions on two different days. One animal and one vehicle property was tested on each day. The same warm-up tasks were administered at the beginning of each session. All subjects imitated at least one of the warm-ups and were praised for doing so.

6. Results

The same pattern of performance found in Experiment 1 was found in the present experiment. Subjects' first choices were entered into an analysis of variance with gender as the between-subject factor and domain (animal, vehicle), exemplar (target, distractor) and assessment (baseline, generalization) as the within-subject factors. Two main effects were found, and one interaction, which is shown in Fig. 3. A main effect for Exemplar, $F(1,14) = 36.61$, $p < .001$, showed that infants demonstrated the properties with the target exemplar ($M = 53\%$) significantly more often that with the distractor exemplar ($M = 14\%$). More properties were demonstrated by the subjects during generalization ($M = 42\%$) than during baseline ($M = 24\%$), $F(1,14) = 19.57$, $p < .001$. As in Experiment 1, an interaction was found between these factors, $F(1,14) = 23.91$, $p < .001$. As can be seen in Fig. 3, performance of the actions significantly increased from baseline to generalization with the target exemplars, $F(1,14) = 23.91$, $p < .001$, (baseline $M =$ 32%, generalization $M = 72\%$), but not with the distractors, $F(1,14) = .08$, (baseline $M = 14\%$, generalization $M = 13\%$). Performance with the targets was significantly greater than with the distractors both at baseline, $F(1,14) = 5.54$, $p < .05$, and generalization, $F(1,14) = 45.95$, $p < .001$. There were no effects of domain or of gender.

An examination of first and second choices was made next. As in Experiment 1, subjects who first demonstrated a property with an inappropriate exemplar frequently made second selections. At baseline, on 5 of the 9 occasions when subjects chose the inappropriate exemplar first, they went on to demonstrate the property with the appropriate exemplar as well.

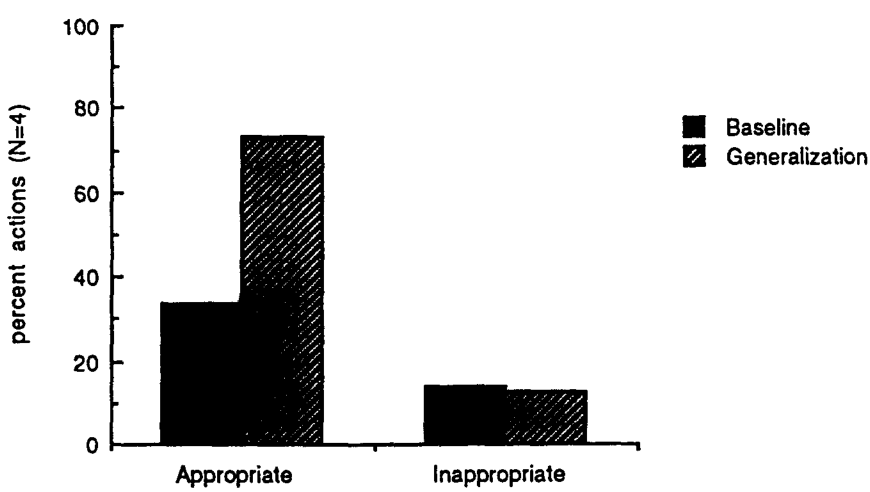

Fig. 3. Experiment 2. Percentage of first actions performed to target and distractor objects.

This also occurred at generalization on 7 out of the 8 inappropriate first choices. However, subjects who demonstrated a property first on an inappropriate exemplar never overgeneralized that property to an inappropriate exemplar at baseline and did so only 10% of the time at generalization. This difference in conditional probabilities was significant, $t(15) = 2.57$, $p < .03$, one-tailed test.

In addition to these analyses, chi-square tests were conducted to see if there were any differences in subjects' demonstrations of the four properties tested and the various exemplars used. No differences were found: subjects used each of the target exemplars to demonstrate the appropriate animal and vehicle properties equally often. Finally, as in Experiment 1, spontaneous demonstrations of the tested properties were noted. Eight subjects spontaneously showed that the ox and armadillo could ride the shoveler, tractor, forklift and crane. Only one inappropriate property was carried out spontaneously: one subject demonstrated giving the cup a ride.

7. Discussion

Experiment 2 not only provided a replication of the findings from Experiment 1 but also increases the certainty that the generalization observed in both experiments was indeed generalization and not merely a reproduction of previously observed events. Fourteen-month-olds would almost certainly not have seen an armadillo sleeping in a bed, or a forklift being started with a key. It seems clear that infants of this age recognize a wide variety of exemplars from the domains of animals and vehicles and have learned behaviors that are appropriate to these domains. Although the behaviors were modeled on and imitated with toy replicas of exemplars from the real domains, the infants were not merely producing play behavior standard to toys, such as scooting model cars or hugging stuffed animals. As another example, they did not put the key to the little models of animals, although wind-up animal toys are common in our culture and it is plausible that some of the infants would have seen such toys being wound up. Instead, the appropriateness of their imitations and their reluctance to cross domain boundaries makes it clear that the toys were being treated as representative of real-world domains. Indeed, in addition to learning through observation of real animals and vehicles acting in the environment, our culture uses games, play scenarios, TV, and picture books to teach children about the properties that we are investigating. We suspect that 14-month-olds are not old enough to have learned a great deal yet from play routines (as opposed to observation of the world), although this is an issue that requires a separate line of investigation.

Regardless of how the relevant knowledge was learned, we could explore its stability and at the same time test whether or not the infants performed as they did merely because they were imitating the games that the adult modeler was playing with the objects. In Experiment 3 we modeled the same actions used in the first two experiments on both appropriate and inappropriate exemplars. Then for the test we gave the infants the exact objects that had been used in the modeling and measured the extent to which the infants would exclude the inappropriate object from their imitations in spite of the experimenter's example. This experiment is similar to the work carried out by Killen and Uzgiris (1981), described earlier, showing reluctance on the part of infants to engage in counterconventional behavior. However, they modeled single actions, such as scooting a car or a cup. We modeled a relational event in which an exemplar of a domain interacted with both an appropriate and an inappropriate object, such as making both a dog and a car drink from a cup. In addition, we also tested recall of these actions. Even if infants proved willing to imitate an action on an inappropriate object immediately following its modeling, they might not remember it as well the following day.

EXPERIMENT 3

8. Method

8.1. Subjects

Sixteen 14-month-olds were tested (mean = 14 months, 16 days; range = 14 months, 2 days to 14 months, 27 days). Preliminary data indicated that presenting subjects with all four tasks and modeling the properties with both the target and distractor objects was too distracting (as shown by order effects); therefore, the subjects were randomly assigned to two groups so that each subject was tested on only one animal and one vehicle property. Eight subjects were assigned to a group that was tested on drinking from a cup and using a key. Another 8 subjects were tested on sleeping in a bed and giving a passenger a ride.

8.2. Procedure

The same properties tested in Experiments 1 and 2 were tested again, using the same props. However, in Experiment 3 the subjects were shown the properties modeled on both the appropriate (target) and inappropriate (distractor) exemplars. The animal exemplars were a dog, cat, rabbit, and bird; the vehicle exemplars were a car, truck, motorcycle, and airplane. These exemplars were the same as used in Experiment 1, except the fish and bus were not used in this experiment (see Fig. 1). Each property was tested with one animal and one vehicle. After baseline was assessed for the relevant objects, the experimenter modeled a property on both an animal and vehicle (e.g., drinking was modeled with a dog and a car, or with a rabbit and a motorcycle). The experimenter would place one exemplar to her right and the other to her left. With the prop in her hand, she would turn to one exemplar, model the property three times, then turn to the other exemplar and model the property another three times. In modeling the property with the inappropriate exemplar, the experimenter would place the prop closest to the part of the exemplar that might be used if the action were appropriate. For example, the car was given a drink by placing the cup beneath the headlights, and the airplane was given a drink by placing the cup under the "nose" of the plane. In one task the experimenter modeled the property with the appropriate exemplar first; in the other task the inappropriate exemplar was used first. Placement of the appropriate and inappropriate exemplar to the experimenter's right or left at modeling occurred equally often. Imitation was assessed immediately after the property was modeled on the two exemplars. The exemplars were simultaneously placed to the right and left of the subject. The experimenter then handed the subject the prop with the accompanying vocalization (e.g., "sip, sip, umm, good").

After this session, subjects returned to the lab 24 hours later. Deferred imitation (recall) of what had been modeled on the previous day was assessed. The experimenter again handed the test items (the prop and two exemplars) to the subject in the same order and manner as had been done the previous day at baseline (i.e., without modeling the actions).

The same two warm-up tasks used in Experiments 1 and 2 were administered at the beginning of the sessions on both days. Imitation of the actions was encouraged and praised. Note that the warm-up tasks encouraged infants to imitate the same action on two perceptually different sets of objects.

9. Results

Even though the properties were modeled on both the target and distractor exemplars, subjects performed in much the same fashion as they did in Experiments 1 and 2. A mixed-design analysis of variance was conducted with group (group 1: tested on drinks, uses with a key; group 2: tested on sleeps, takes a rider) as the between-subject factor and exemplar (appropriate, inappropriate) and assessment (baseline, imitation, recall) as the within-subject factors. Subjects' first choices were entered as the dependent measure. Main effects for exemplar and assessment were found. The subjects demonstrated the properties with appropriate exemplars ($M =$ 46%) significantly more often than with inappropriate exemplars ($M =$ 21%), $F(1, 14) = 8.13$, $p < .05$. A main effect was also found for assessment, $F(2, 13) = 4.74$, $p < .05$. Tukey tests ($p = .05$) showed that the

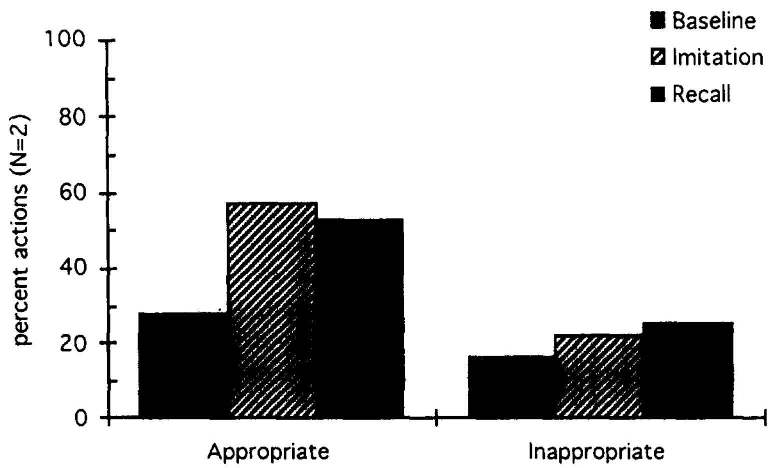

Fig. 4. Experiment 3. Percentage of first actions performed to appropriate and inappropriate objects.

subjects imitated ($M = 39\%$) and recalled ($M = 39\%$) more of the properties than they spontaneously performed at baseline ($M = 22\%$).

As can be seen in Fig. 4, performance increased for both the appropriate and inappropriate exemplars after the properties were modeled. Because of this increase with both exemplars, an interaction involving exemplar and assessment did not approach significance, $F(2, 13) = 1.51$, $p > .10$. Nevertheless, we thought it informative to conduct post hoc analyses on the appropriate exemplar data separately from the inappropriate exemplar data. These analyses indicated that performance increased significantly over the assessment periods for the appropriate exemplars, $F(2, 13) = 4.68$, $p > .05$, (baseline $M = 28\%$, imitation $M = 57\%$, recall $M = 53\%$). However, the increase was not significant for the inappropriate exemplars, $F(2, 13) = 0.88$, (baseline $M = 16\%$, imitation $M = 22\%$, recall $M = 25\%$). Thus, the findings are consistent with the generalization data of Experiments 1 and 2. It may also be noted that there was no drop-off in performance of actions on the inappropriate exemplars after a 24-hour delay; performance of these actions remained at a relatively low level throughout.

An examination of the pattern of first and second choices showed that second selections were made more frequently to the inappropriate exemplars in the present experiment than in Experiments 1 and 2. In Experiment 1, inappropriate second selections were made on average 14% of the time and in Experiment 2, they were made 5% of the time. That is, subjects only infrequently overgeneralized a property to a distractor after having correctly generalized the property to an appropriate exemplar. However, in the present experiment, after making appropriate selections for their first choices, subjects then went on to make inappropriate second selections on average 46% of the time (imitation = 56%; recall = 35%). This was equally true whether modeling with the appropriate or inappropriate exemplar occurred first. This finding indicates that the subjects did attend to the experimenter when the properties were modeled with inappropriate objects; however, they still more frequently chose the appropriate objects for their first imitation or recall of the property.

The data were inspected for differences in performance of the properties tested or the exemplars used. No significant differences were found; subjects did as well on the occasions when the two most similar items (the bird and the airplane, shown in Fig. 1) were paired as when any other combinations of animal and vehicle were used. In addition, there was no drop-off in performance with the airplane when giving a ride in this experiment, suggesting that the earlier finding in Experiment 1 was not reliable. There were fewer spontaneous demonstrations in this study than found in the previous two experiments. Five subjects showed an animal riding on a vehicle, with one subject giving the bird a ride on the airplane!

GENERAL DISCUSSION

In three experiments we studied the inductive generalizations that 14-month-old infants have made about the animal and vehicle domains. When the infants were shown appropriate actions performed on a model of an animal (or a vehicle), they were willing to imitate these actions on any other exemplars from the same domain no matter how perceptually dissimilar these were from the object used in the modeling. The infants were less willing to imitate these actions on exemplars from a different domain. This was so even when the exemplar from the inappropriate domain was highly similar in appearance to the appropriate exemplar, as in Experiment 3 when a bird and airplane were pitted against each other. Furthermore, the infants were reluctant to cross domain boundaries in their imitations even when the actions on inappropriate objects were modeled for them. This is an important finding, because it shows that the generalizations the infants have made are powerful ones; they do not adopt a simple strategy of assuming any property modeled by an adult is appropriate. It is also of interest that no differences were found in the likelihood of making generalizations in the two domains studied, one of which consisted of natural kinds (animals), the other of which consisted of artifacts (vehicles).

We assume that the generalizations our subjects demonstrated were implicit in the sense that the infants were not deliberately making inferences nor answering questions about them. Nevertheless, the data tell us that inductive inferences do not require labeling or the use of language, and that 14-month-old infants, just like older children, not only make inferences but constrain their inductive generalizations on the basis of their conceptual categories. Since at this young age conceptual categories tend to be rather general (Mandler et al., 1991; Mandler & McDonough, 1993), the generalization that occurs is widespread; indeed, it appears to be domain-wide for animals and vehicles, constrained only by the boundaries of the domains themselves. The data showed no influence of perceptual similarity between modeled and test exemplars on likelihood of generalization. The infants were just as likely to generalize from a dog to a fish, bird, or an armadillo as to a cat or a rabbit. Similarly, they were as likely to generalize from a car to an airplane, motorcycle, or forklift as to a truck or a bus. This conclusion follows, of course, from the previous one: the generalizations were uniform across the entire domain. Thus, we found no evidence that infants make inductions on the basis of responsivity to uninterpreted animal similarity, as speculated by Quine (1977), or on the basis of original sim, as speculated by Keil (1991).

Instead, the present data agree with research on older children, showing that their inductions are also more influenced by conceptual category than by perceptual similarity (e.g., Gelman & Markman, 1986). The main difference between our data on 14-month-olds and those of older children seems to be in the size of the "smallest" classes that constrain the inductions. By 4 years of age preschool children's smallest classes are already at the subordinate level, and so they make the most inductive inferences at that level (Gelman, 1988). Our previous categorization data (Mandler et al., 1991; Mandler & McDonough, 1993) indicate that relatively little conceptual subdivision of the animal and vehicle domains has yet taken place at 14 months. Insofar as one can use willingness to generalize as a measure of the breadth of a class, the uniformity of the generalization the 14-month-olds showed is another indication of the lack of subdivision. Indeed, the very breadth of their domain-level categories led the infants to make some overgeneralizations. For example, making fish drink is a consequence of having made a domain-wide generalization, since it is unlikely the infants would have observed it, but is also a conclusion that will eventually have to be revised. A similar comment might be made about "opening" a forklift or an airplane with a key.

Many questions remain unanswered by this first foray into the study of inductive generalization in infancy. First, although we have shown domain-wide generalization uninfluenced by similarity, it could be that if infants were given several choices from a single domain, similarity might then have some effect. For example, if a property is modeled using a dog, infants might be more likely to choose another dog than a cat or a bird for their imitation of the property. We are currently testing this proposition; we have not yet tested birds, but to date we find that 14-month-olds are as likely to choose a cat or a rabbit as another dog.

Second, we do not yet know whether there are constraints on the kinds of properties that infants will generalize. By the time children are 3 years of age, they know that properties such as dirtiness, ownership, age, and so forth, are not generalizable (Gelman, 1988). Unless they have older siblings, 14-month-olds may not know some of these properties (e.g., ownership), and in other cases (e.g., dirtiness) we simply do not know whether they would restrict their applicability to certain classes or not. Experiments to test infants' sensitivity to generalizable versus nongeneralizable properties are in progress.

Third, we assume that our 14-month-old subjects had already made the generalizations that were tested in these experiments. We do not suppose that we taught them that armadillos drink or that airplanes are "keyed". Instead we assume that our procedures merely enabled the infants to demonstrate inferences about whole domains that they had already made, such as that animals drink. We are currently extending this work to 9- and 11-month olds, on the assumption that at some early age we will find subjects who have not yet made the relevant generalizations. In our work on these ages to date, however, we find that at both these ages, although more clearly at 11 months, the infants show effects similar to those found in the present experiments.

Finally, we have yet to test infants' knowledge about so-called basic-level concept properties. In the present experiments we only tested characteristics generally appropriate for a large domain: drinking and sleeping for animals and using a key and transporting people for vehicles. It is possible that by 14 months infants would restrict some kinds of properties, such as barking, to dogs, and refuse to generalize them to other subclasses such as cats. Indeed, it seems likely that it is the very process of learning properties (other than perceptual appearance) specific to subclasses of animals and vehicles that makes the earlier formed perceptual subdivisions meaningful. In that case, the process of inducing subclass properties would be intimately bound up with the formation of the smaller conceptual classes that will eventually come to dominate the induction process.

In conclusion, this account of early processes of induction has called upon a distinction between perceptual and conceptual categories. Indeed, this is the third task in which we have found it necessary to distinguish between these two kinds of categorization. The first task is the object-manipulation task in which sequential touching of objects is used (Mandler & Bauer, 1988; Mandler et al., 1991). This task, which is the closest approximation to classification that can be used with 1- to 2-year-olds (Sugarman, 1983), shows that global categorization of whole domains, whose exemplars do not look much alike, occurs in the absence of most basic-level categorization, in which exemplars do look alike. Immediately one must ask what information is being used when infants categorize exemplars such as animals that are not perceptually alike and fail to categorize those that are. The second task, used with younger infants, is the object-examination task. This task is somewhat further removed from classification in that it is a kind of habituation–dishabituation task. Yet this task produces the same kind of results as the sequential touching task, with global categorization occurring more often that categorization based upon overall perceptual similarity. The present experiments used still another task, namely imitation. Once again, responsivity to global domains, accompanied by a lack of responsivity to the more perceptually similar subclasses of these domains, was found.

It is clear from these various results that infants are using some other basis than simple perceptual similarity for at least some of their categorization. One might try to explain the results by assuming that dissimilarity between classes is more important than similarity within classes. When contrasting classes are taken from different domains, the between-class contrast is often perceptually greater than when the classes are taken from the same domain. Nevertheless, that is not always the case, and infants easily distinguish between models of birds and airplanes (Mandler & McDonough, 1993). Indeed, inspection of the bird and airplane in Figure 1 suggests that there are at least some cases in which the physical similarity of a between-domain contrast is less than that of a within-domain contrast.

Alternatively, one might try to explain the results by assuming that whatever the features are that allow identification of animals or vehicles as members of their respective domains, they are so perceptually salient that they are the first to be discriminated by infants. In this approach, one would deny that the first categories to be formed are basic-level ones; instead, one would assert that all categorization in infancy is perceptual in nature, but that the first categories formed by this process are global, with the more differentiated basic-level categories being a later development. However, the data of Quinn and Eimas and their colleagues indicate that very young babies *are* responsive to the kind of detailed perceptual similarity associated with what are usually called basic-level categories; indeed these may be an even earlier accomplishment than the domain-level categorization we have shown, since 3-month-olds categorize pictures of horses as different from pictures of zebras (Eimas & Quinn, 1994).

The most parsimonious solution to this confusing developmental picture, in which skills appear to be gained and then lost, is that more than one kind of categorization is taking place. One kind relies on perceptual schematizing or prototyping and begins very early in infancy. It produces categorization of objects such as dogs and cats, and this can be accomplished strictly as a response to the similarity of patterns, even when no meaning has yet been associated with these patterns. This kind of perceptual categorization provides stability and familiarity in the infant's world, but in itself does not provide a basis for understanding what kinds of things these categorized objects are. That requires a different sort of process. We have suggested that at the same time that infants are forming perceptual categories they are also beginning to *analyze* perceptual displays in such a way as to give them meaning (Mandler, 1992). Meaning also requires noticing similarities among things, but similarities that have a different character than the commonalities found in perceptual patterns. The notion of kind (i.e., being the same thing as opposed to merely looking similar) is a more abstract characterization of an object than its usual perceptual description; it arises from analysis of the information that broadly characterizes entire domains of objects. For example, different animals begin motion in very different ways (moving limbs, opening wings, wriggling) but all start to move by themselves, whereas vehicles do not. Descriptions such as self-moving are derived from a more abstract kind of analysis of objects, and constitute their first conceptual, as opposed to perceptual, descriptions (Mandler, 1992). These descriptions or meanings are crude, but they form the core notions of kinds of things, and therefore create the conceptual categories that allow inductive inferences to be made.

No one would expect adults to make inductive inferences solely on the basis of perceptual similarity, ignoring conceptual class membership. Yet babies have been assumed to do exactly that. Of course, adults' perceptual processes are suffused with conceptual information, and therefore adults are almost always using both perceptual and conceptual information in their categorization. It is perhaps only in infancy that it is even possible to separate purely perceptual processes from meaning-based ones. It is interesting, therefore, that babies who in principle have purely perceptual categories available to them to use for purposes of making inductive inferences, nevertheless do not do so. In this respect they are less influenced by perceptual factors than adults (a perhaps surprising reversal of the common saying that young children are more influenced by perceptual appearances than are adults). Nevertheless, like adults they use perceptual information to tell them about conceptual category membership, and when it comes to making inferences as to shared characteristics, it is that conceptual membership which matters.

Acknowledgements

Preparation of this paper was supported in part by NSF research grant DBS92-21867. Thanks to Natasha Brown, Jeannie Chufo, and Jennifer Swerdlow for their assistance in data collection and analysis.

References

Baldwin, D.A., Markman, E.M., & Melartin, R.L. (1993). Infants' ability to draw inferences about nonobvious object properties: evidence from exploratory play. *Cognitive Development*, *64*, 711–728.

Bauer, P.J., & Dow, G.A. (1994). Episodic memory in 16- and 20-month-old children: specifics are generalized but not forgotten. *Developmental Psychology*, *30*, 403–417.

Carey, S. (1985). *Conceptual change in childhood.* Cambridge, MA: MIT Press.

Clark, E.V. (1983). Meanings and concepts. In J.H. Flavell & E.M. Markman (Vol. Eds.), P.H. Mussen (Series Ed.), *Handbook of child psychology. Vol. 3. Cognitive development.* New York: Wiley.

Eimas, P.D. (1994). Categorization in early infancy and the continuity of development. *Cognition*, *50*, 83–93.

Eimas, P.D., & Quinn, P.C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, *65*, 903–917.

Gelman, S.A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology*, *20*, 65–95.

Gelman, S.A., & Coley, J.D. (1990). The importance of knowing a dodo is a bird: categories and inferences in 2-year-old children. *Developmental Psychology*, *26*, 796–804.

Gelman, S.A., & Markman, E.M. (1986). Categories and induction in young children. *Cognition*, *23*, 183–209.

Gelman, S.A., & Markman, E.M. (1987). Young children's inductions from natural kinds: the role of categories and appearances. *Child Development*, *58*, 1532–1541.

Gelman, S.A., & O'Reilley, A.W. (1988). Children's inductive inferences within superordinate categories: the role of language and category structure. *Child Development*, *59*, 876–887.

Keil, F.C. (1991). The emergence of theoretical beliefs as constraints on concepts. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind* (pp. 237–256). Hillsdale, NJ: Erlbaum.

Killen, M., & Uzgiris, I.C. (1981). Imitation of actions with objects: the role of social meaning. *Journal of Genetic Psychology*, *138*, 219–229.

Mandler, J.M. (1992). How to build a baby: II. Conceptual primitives. *Psychological Review*, *99*, 587–604.

Mandler, J.M. & Bauer, P.J. (1988). The cradle of categorization: Is the basic level basic? *Cognitive Development*, *3*, 247–264.

Mandler, J.M., Bauer, P.J., & McDonough, L. (1991). Separating the sheep from the goats: differentiating global categories. *Cognitive Psychology*, *23*, 263–298.

Mandler, J.M., & McDonough, L. (1993). Concept formation in infancy. *Cognitive Development*, *8*, 291–318.

Mervis, C.B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*, 89–115.

Piaget, J. (1951). *Play, dreams, and imitation in childhood*. London: Kegan Paul, Trench, & Trubner.

Quine, W.V. (1977). Natural kinds. In S.P. Schwartz (Ed.), *Naming, necessity, and natural kinds*. (pp. 155–177). Ithaca, NY: Cornell University Press.

Quinn, P.C., & Eimas, P.D. (1986). On categorization in early infancy. *Merrill-Palmer Quarterly*, *32*, 331–363.

Quinn, P.C., & Eimas, P.D. (1994). *Young infants' use of facial configuration information permits differentiation of natural animal species*. Paper presented at the meetings of the International Society for Infant Studies, Paris.

Quinn, P.C., Eimas, P.D., & Rosenkrantz, S.L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old infants. *Perception*, *22*, 463–475.

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

Rosch, E., Mervis, C.B., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *3*, 382–439.

Sugarman, S. (1983). *Children's early thought*. Cambridge, UK: Cambridge University Press.