Children’s Use of Generics in Inductive Inferences

Susan A. Gelman, Jon R. Star, and Jonathan E. Flukes

Department of Psychology
University of Michigan

What conditions foster or constrain children’s category-based induction? This study examined the role of language, specifically the scope of noun phrases used to convey novel property information. We focused on generic noun phrases, which are especially common in child-directed speech and have been argued to play an important role in the growth of category knowledge. In Study 1, 4-year-olds and adults were taught novel properties about familiar categories, using 1 of 3 types of wording: a generic noun phrase (e.g., “Bears like to eat ants”), a universally quantified noun phrase (e.g., “All bears like to eat ants”), or an indefinite plural noun phrase (e.g., “Some bears like to eat ants”). Study 2 was a follow-up study with adults using a more sensitive task. Results indicated sensitivity to type of wording among both preschoolers and adults, with “all” eliciting the most inferences, “some” eliciting the fewest inferences, and generics in between “all” and “some.” However, children made fewer category-based inferences from generics than did adults. Together, these studies suggest that (a) preschool children as well as adults distinguish generics from both “all” and “some,” and (b) age differences exist in the interpretation of generic noun phrases.

Much of human reasoning depends on our ability to draw inferences beyond the evidence immediately available. This inductive process is often based on the use of categories (Gelman & Medin, 1993; Osherson, Smith, Wilkie, Lopez, & Shafir, 1990). Thus, categories are crucial to the organization and extension of our knowledge about the world (Smith & Medin, 1981). We expect members of a category to share novel features, such as external appearance (e.g., all dogs have four legs), chemical structure (e.g., all cells contain nuclei), internal structure (e.g., all bears have the same number of bones), certain behaviors (e.g., all fish breathe through gills), and other nonobvious properties.

Requests for reprints should be sent to Susan Gelman, Department of Psychology, University of Michigan, 525 E. University Ave., Ann Arbor MI 48109-1109. E-mail: gelman@umich.edu
Much evidence exists that even young children recognize the importance of category membership for drawing novel inferences. Preschool children expect category members to share nonobvious similarities, even in the presence of external dissimilarities, and they judge internal, nonvisible parts as important to the identity of novel items (Déak & Bauer, 1996; Gelman & Coley, 1990; Gelman & Markman, 1986; Keil, Smith, Simons, & Levin, 1998). Indeed, given the wealth of potential unobservable similarity between members of a category, facts learned about a single category member often generalize to the entire category.

However, children (and adults) cannot always assume that newly learned properties generalize to an entire category for at least two reasons. First, each object is a member of a varied set of categories (e.g., the same object is at once a cat, a pet, a mammal, and a vertebrate), thus raising the question of how one selects the category to which a property applies. For example, after observing the body temperature of your pet cat, should you generalize it to cats, pets, mammals, or vertebrates? Second, some properties are idiosyncratic and thus do not warrant generalization along the lines of conventional categories at all (e.g., “fell on the floor this morning”; Gelman, 1988). Thus, the human capacity to generalize raises the question of how this capacity is constrained (Goodman, 1973).

It is likely that multiple potential sources of information help guide and constrain inductive inferences. First, the human conceptual system may itself be fundamentally limited such that certain sorts of inductive inferences never get considered (see Peirce, 1992). Thus, on learning the body temperature of your pet cat, you do not attempt to generalize this fact to objects only thematically related to her (e.g., her food bowl or favorite toy; see Markman, 1992, regarding the limited inductive potential of thematic categories). Second, given enough direct experience, statistical information may prove important (Holland, Holyoak, Nisbett, & Thagard, 1986; Fong, Krantz, & Nisbett, 1986). For example, one could make a study of the body temperatures of a range of cats and other animals, and use information from this sample to help infer whether body temperature generalizes beyond that of your own pet. This sort of statistical reasoning is often used in the sciences. Third, content information and folk theories presumably help determine the scope of a novel generalization. For example, you could use your knowledge of the causal links among body size, metabolism, and temperature to draw reasonable inferences regarding the relevance of your cat’s temperature for that of other animals.

Fourth, and of particular interest here, linguistic information may play an important role in guiding or constraining inductive inferences—particularly for children, given their limited first-hand experience, limited content knowledge, and apparent difficulty reasoning from statistical evidence. Often, novel properties are encountered in a social context, including language that frames the novel property in some way. One clue that language can provide derives from the choice of noun a speaker uses to refer to an object (e.g., cat, mammal, vertebrate; see Callanan,
1990, for interesting evidence regarding the systematicity with which parent conversations regarding everyday objects reflect different sorts of properties for categories at different levels of abstraction). A second clue in language comes from the forms of noun phrase, including choice of quantifiers (e.g., all, some, and many), used to express the speaker’s belief regarding the scope of a proposition. Thus, “The body temperature of my cat is X degrees” has a very different implication from “The body temperature of all cats is X degrees” or “The body temperature of some cats is X degrees.” Indeed, without language, it is not clear how one could convey the scope of a proposition to others.

Thus, language would seem to provide a rich source of information to children concerning the potential scope of inductive inferences. In this study, we examine whether and how children make use of information in linguistic quantifiers to direct their inductions. We focus specifically on three types of noun phrases: (a) generics (e.g., “Dogs have four legs”), (b) universal quantifiers (e.g., “All bears have claws”), and (c) indefinite plurals (e.g., “Some turtles have spotted shells”). Each of these language forms differs semantically, as discussed next.

**GENERIC NOUN PHRASES**

Generic expressions are used to express properties that are law-like or definitional (Carlson & Pelletier, 1995; Dahl, 1975). For example, the statement “Dogs have four legs” means that, under normal circumstances, each member of the category “dogs” has the property of possessing four legs. Although a generic statement applies broadly to a category, it is not considered false by the presence of individual category members for whom the property does not apply (e.g., a dog who has lost a leg in an accident). In addition, generics may not even be true for a majority of cat-

---

1 For the purposes of this article, we consider inductive inferences to be those that involve a degree of uncertainty, in contrast to deductive inferences in which conclusions follow necessarily and logically from stated premises (see Skyrms, 1975, for this use of “induction”). Note that this definition differs from the view that induction involves reasoning from particulars to generalities. In contrast, we treat as “inductive” any inference that cannot be fully supported by logical principles and thus can include reasoning from particulars to particulars (e.g., If this bear likes to eat flies, does that bear like to eat flies?), from generalities to particulars (e.g., If bears like to eat flies, does this bear like to eat flies?), from generalities to generalities (e.g., If bears like to eat flies, do all bears like to eat flies?), or from particulars to generalities (e.g., If this bear likes to eat flies, do bears like to eat flies?). Inferences from generic statements and from “some” statements clearly qualify as “inductive,” using the present definition, because both generic statements and “some” statements lack the force of logical necessity when reasoning about category instances. For example, based on the statement that “bears like to eat flies,” we cannot deduce with certainty that any particular bear likes to eat flies. However, inferences from “all” statements (e.g., “all bears like to eat flies”) are more properly considered deductive rather than inductive, as long as “all” refers to the entire category. Note that, because it is not yet clear precisely how children interpret “all,” for our purposes we consider inferences from universal quantifiers to be potentially inductive as well.
egory members (McCawley, 1981). For example, the generic statement “Birds lay eggs” persists, despite the fact that it applies to less than half of the category of birds (excluding male birds and infants; Gelman, Coley, Rosengren, Hartman, & Pappas, 1998; Prasada, 2000). As a result, facts stated generically may be particularly robust against counterevidence (e.g., “birds fly” allows for penguins, but “all birds fly” does not).

Generic noun phrases are of particular interest in children’s inferences, as they are common in child-directed speech. In a study of maternal speech to 35-month-olds, Gelman et al. (1998) found that 27 out of 30 mothers used at least one generic noun phrase during a brief book-reading session (approximately 15-min long) and that generic noun phrases exceeded 5% of maternal speech. In a similar study, Pappas and Gelman (1998) found that generic noun phrases appeared in both maternal and young children’s speech. Eleven percent of maternal utterances on this task included a generic; in addition, more than half of the 2- to 4-year-old children produced at least one generic noun phrase during brief picture book reading sessions.²

Moreover, generics may foster category-based inferences. Shipley (1993) proposed that a grouping becomes a “kind” (i.e., richly structured and inference promoting) when a person “projects a property onto individual members of a class” (p. 270). In other words, when a person learns a new property of a category as a whole, that category is hypothesized to cohere in novel ways. For example, if a child hears that “Birds fly,” he or she may hypothesize that one can make a variety of other novel inferences about the class of birds. The making of inductive inferences about a generic class may thus help in the transformation of a class to a kind. Prasada (2000) similarly suggested, “It is likely that one of the primary mechanisms by which children acquire generic knowledge is through language” (p. 70).

Although several studies document the frequency of generics in parental speech, and to some extent even in children’s speech, to date there has been no systematic study of generic interpretation in children. As Prasada (2000) noted, “Unfortunately, there are as yet no studies that have directly investigated children’s in-

²Although generics occur in only a small percentage of parental speech, this frequency represents a substantial amount of input to children. Nouns can function in many different ways besides generic reference, including singular definite reference, general definite reference, nonreferring definite reference, distributive general reference, collective general reference, specific indefinite reference, and non-specific indefinite reference (Lyons, 1977, pp. 177–197). Accordingly, even the most salient of noun phrase types will occur in less than the majority of utterances. To determine the relative salience of generics, it is therefore misleading to consider the proportion of speech containing generics, and more meaningful to consider the absolute frequency of such speech. The rate of generics in maternal speech is comparable to the rate that mothers produce causal language (Hickling & Wellman, 2001) and exceeds the rate that children produce genuine psychological references to thoughts and beliefs at 6 years (Bartsch & Wellman, 1995). In our own sample, the rate of generic use was greater than the rate at which mothers talked about object size (3.09% of utterances), color (1.96% of utterances), number (0.77% of utterances), shape (0.35% of utterances), or texture (0.22% of utterances; Gelman, 2000).
terpretations of generic statements and thus it is difficult to know the extent to
which children use generic statements to acquire generic knowledge” (p. 70).

UNIVERSAL QUANTIFIERS

For adults, universal quantifiers (e.g., “All bears have claws”) refer to an entire
category, without exception. Such statements are the most inclusive and unambiguous in
indicating the inductive potential of a category, and yet they are easily falsified by a
single counterexample. Although traditionally children were thought not to under-
stand logical quantification before the stage of concrete operations (Inhelder &
Piaget, 1964), simplified tasks have yielded above-chance performance in children
as young as 4 years of age (Smith, 1979, 1980). For example, Smith (1979, 1980)
found that 4-year-olds were able to do the following: correctly identify appropriate
referents for “all” (e.g., “Are all of the magnets on the box?”), correctly answer inclusion
questions (e.g., of the form, “Are all Xs Ys?”), and successfully complete property
inference tasks (e.g., given that “All X’s have _____” and that Y is either subordinate
or superordinate to X, answering the question, “Do all Ys have _____?”). In
addition, Smith (1979) found that children were quite articulate in justifying their
correct answers in each of the above types of tasks.

These data do not imply that children understand “all” in precisely the same
manner as adults. Indeed, children continue to make errors in their interpretation of
“all” regarding subtle aspects of meaning (e.g., Brooks & Braine, 1996; Hanlon,
1987). However, for the purposes of this discussion, the relevant point is that pre-
school-aged children appear to understand the basic semantic implications of “all,”
at least on some tasks. This finding suggests that logical quantifiers appearing in
adult speech could be effective in transmitting information regarding inductive poten-
tial to children.

Nonetheless, statements using “all” as a universal quantifier are quite rare in par-
ents’ speech. In the study of maternal speech to 35-month-olds, mentioned earlier,
Gelman et al. (1998) found that only 2 out of 30 mothers used “all” as a universal
quantifier, during a 15-min book-reading session. In fact, out of 120 “all” statements
produced during these sessions, only 2 were universal quantifier uses; the remaining
uses were either for nonquantification purposes (e.g., “Are you all done using that?”)
or to refer to a specific context (e.g., “All the puzzle pieces go back in this box”).

INDEFINITE PLURALS

Indefinite plurals (e.g., “some”) refer to a specific subset of category members.
The work examining “all” found that children are likewise competent at under-
standing “some” (Smith, 1979, 1980). Such work has established that statements
such as “Some Xs are Ys” are deemed appropriate by children when Xs include
both Ys and non-Ys. Little is known, however, regarding the implications of “some” for inductive reasoning. Specifically, it is not yet known how broadly children would generalize a new fact that is predicated of “some” category instances (e.g., “Some dogs have leukocytes in their blood”). Of particular interest is how such statements would compare to generics, in inductive potential.

THESE STUDIES

These studies examine whether preschool children can make use of language cues (generics, “all,” and “some”) to guide their inductive inferences. On each of a series of trials, participants were taught a novel property about an exemplar animal using one of three language forms (generic noun phrase, universally quantified noun phrase [using “all”], or indefinite plural noun phrase [using “some”]). We measured how frequently they generalized the novel property to novel instances of the category. We selected the youngest age at which children have been shown to understand “all” and “some” on deductive inference tasks (see Smith, 1979, 1980). Adults were also included to measure the developmental endpoint in the semantics of these expressions. We predicted that, for adults, “all” statements would generate the most extensive inferences, followed by generic statements, followed by “some” statements.

Study 1 was the primary study, comparing directly generics, “all,” and “some” in 4-year-olds and adults. We also varied the typicality of the test pictures, and the order in which linguistic cues were provided to participants (namely, before or after the set of response choices). Typicality was found to play a role in children’s category induction in prior work (see Lopez, Gelman, Gutheil, & Smith, 1992). Inclusion of less typical exemplars thus provides a strong test of the power of language for guiding inductive inferences. Whereas we might expect children to generalize novel properties to other typical category exemplars even in the absence of language, it is of interest to determine what language forms enable children to make less obvious inferences. The order of linguistic cues enables us to examine information-processing effects on these inferences.

Study 2 focused specifically on adults’ interpretation of generics, and used a more sensitive task to determine more precisely how adults interpret generics relative to “all” and “some.”

STUDY 1

Method

Participants

Thirty-seven preschool children (M = 4 years; 7 months, range = 3;11–5;5, 23 girls and 14 boys) and 36 adults (23 women and 13 men) participated in the main
study. The children were predominantly White, middle-class residents of a medium-sized midwestern city, attending preschool. Approximately one half of the adults were enrolled in an Introductory Psychology class at the University of Michigan and participated in this experiment for course credit. The remainder were volunteers who received a gift certificate to a local ice-cream shop in exchange for participating. An additional 10 preschoolers and 18 undergraduates participated in pretesting of the materials.

**Materials**

Materials for the children’s study included a picture book that contained 9 sets of picture cards (6 picture cards per set, for a total of 54 pictures) and 9 question pages (1 for each set of picture cards). In addition, there was one clue book and two response cards.

Each question page contained a question of the form, “Which ones _______?” For example, for the category “bears,” the question page said, “Which ones like to eat ants?”

Each set of picture cards included three highly typical instances of the target animal category and three moderately typical instances of that category. Each animal picture card contained a 2- x 3-in. (5.1- x 7.6-cm) color photograph or realistic color drawing. The six picture cards were arranged in two rows of three and were placed on the page in a fixed, random position. Across sets, position of the highly versus moderately typical pictures varied systematically, with the constraint that all three highly typical pictures never appeared in the same vertical row, and all three moderately typical pictures never appeared in the same vertical row. We also avoided placing all three pictures of the same typicality level (either highly or moderately typical) in a cluster at one end of the page (e.g., two highly typical pictures in the left-most column, with a third highly typical picture in the middle column). These constraints on placement were designed to minimize the possibility of confounding typicality and position on the page.

The clue book included nine pages, each containing one property for each animal category (see Table 1). The clues were of the form: (Quantifier) (Animal category) (Property). For example, in the “all” condition for the category “bear,” the clue said, “All bears like to eat ants.”

The two response cards were each 11 x 14 in. (27.9- x 35.5-cm). One response card was used for “yes” responses and the other for “no” responses; the words “yes” and “no” were printed at the top of the respective cards.

**Selection—Validation of Materials**

**Typicality ratings.** We gathered typicality judgments from a separate group of adults (N = 18) to validate the typicality of the animals used in the picture cards. For each animal, six pictures were presented on a single piece of paper in a fixed arrange-
ment. Participants were asked to “judge how good an example of a category various instances of the category are” on a 7-point scale ranging from 1 (a very good example) to 7 (a very poor example). Participants were encouraged to make an intuitive judgment of the overall typicality of each picture, rather than spending a great deal of time trying to figure out their answer. Examples of a very typical chair (a desk chair) and a very atypical chair (a dentist’s chair) were provided. Pictures used in the study were selected as either “highly typical” (M = 1.38, range = 1.0–2.0) or “moderately typical” (M = 3.92, range = 2.6–6.7). Note that we tried to avoid highly atypical pictures, so that all items would be accepted as category members by preschoolers.

**Picture identification.** To ensure that young children could identify all pictures as being members of their respective animal categories, a small study was conducted with a separate group of 4-year-olds (N = 10, M age = 4–8, range = 4;1–5;0). Children were individually shown displays, each containing nine pictures—the six target pictures plus three distractors (e.g., a walrus, giraffe, and block)—and were asked, “Look carefully at the pictures on this page and point to all of the Xs (e.g., bears).” They were then prompted with, “Are there any more or did you get them all?” On this measure, children selected the target pictures 95.2% overall (99% of the typical pictures and 92% of the moderately typical pictures), and never selected the distractor pictures. For each of the nine animal categories, the mean percentage of target pictures chosen ranged from 90% to 100%. Together these responses indicate that preschool children readily accept both the typical and the moderately typical pictures as belonging to their respective categories.

**Procedure**

Each child participant was tested individually by a male researcher. Children were tested in a quiet corner of their class or in a private room and sat across from the researcher on the floor.
"All–some" pretest. Before the main task began, a short pretest was administered to determine if children could distinguish the meanings of the words "all" and "some" (after Smith, 1980). The purpose of the pretest was to remind children of the meanings of these words and to ensure that all participants had the minimal linguistic knowledge required to complete the task. The researcher placed a small, clear plastic cup and four pennies in front of the child, saying, "First we are going to play a little game. Here I have pennies. And here I have a cup. Now watch." A number of pennies (zero, two, three, or four) were placed in the cup, with the remaining pennies left visible outside the cup, to the side. The child was then asked two questions: "Are all of the pennies in the cup?" and "Are some of the pennies in the cup?" (with order of the two questions constant across trials, but counterbalanced across children). Each child received four trials of this pretest, using zero, two, three, and four pennies. Correct performance entailed saying "yes" to the "all" question only when four pennies were in the cup, and saying "yes" to the "some" question only when two or three pennies were in the cup. Only children whose responses fully conformed to this pattern proceeded to the main task.

Introduction to main task. At the conclusion of the pretest, the picture book was placed on the floor facing the child, with the response cards on either side of the picture book. The clue book was placed off to one side. The researcher stated the following:

We are going to play a game here with some picture books. I have two books: a big book with questions and pictures in it, and a little book with clues in it. This is the question book; it has a question on each page. The questions are kind of hard, and you're going to need some help answering them. So that's why we have this book over here; it's the clue book. The clue book tells you things you need to know, to help you answer the questions in the question book. Do you understand? Can you show me the question book? Can you show me the clue book? Great!

Children were randomly assigned to one of two conditions: the clues-first order condition or the clues-second order condition. That is, for one half of the participants, we provided the property information ("clue") first, followed by the test question; for the remainder, we posed the test question first, and then gave the property information in the form of a clue. We varied this factor to examine the effects of differing information-processing demands on performance. Specifically, we hypothesized that children might be better able to make use of the linguistic information when it appeared second, as they would be most receptive to any relevant information that could help answer the test question. In contrast, when the linguistic information appeared first, we hypothesized that children would be less
attentive to the wording and so less likely to make use of that information when answering the question that followed.

**Practice trials.** At this point, four practice sets were given to each child to familiarize him or her with the study procedure. In particular, we wished to make sure that children realized that (a) they should listen carefully to the researcher’s instructions, and (b) they could pick more than one picture in response to each question. For each practice set, the child was shown and read a question (e.g., “What does James want?”), either preceded by a clue (clues-first condition) or followed by a clue (clues-second condition). The clues varied in mentioning either one or two instances of a category. For example, on this item the clue was, “James wants two cars.” The child was then shown two different pictures of the nonanimal category (e.g., of two different cars) and was asked to place each picture, one at a time, on either the “yes” response card or the “no” response card. If a child answered a practice item incorrectly, the researcher repeated the question and the clue, placed the cards in their correct location, and gave an explanation (e.g., “James wants two cars so you put them here where it says yes because James wants two cars.”). No child was excluded on the basis of his or her performance on the practice trials. However, all children answered at least three out of the four practice questions correctly.

Each child was given four further practice sets throughout the session: two practice sets after the third animal trial and another two practice sets after the sixth animal trial. The purpose of these practice sets was to remind children to listen carefully to the experimenter’s words, and to give children some “easy” trials to maintain a high degree of confidence and interest in the task. Responses to these practice sets were not recorded and were not used further.

The various pretests and practice trials are summarized in Table 2.

**Primary task.** The procedure for the clues-first condition was as follows. After the practice sets, the researcher turned to the clue book and said, “Here is the first thing we are going to learn,” and read the property (e.g., “All bears like to eat ants”). He repeated the property and asked the child to repeat the property (to ensure learning of the clue). Then he turned to the relevant page in the picture book, told the child which type of animal they were seeing (e.g., “Here is a page of bears”), and read the question (e.g., “Which ones like to eat ants?”). The facing page displayed the picture of the six animals (in this case, bears).

---

3We opted not to include distractor examples in the main study for two primary reasons: (a) to reduce information-processing demands (i.e., inclusion of distractor pictures would require children to scan nine pictures at once while holding in mind the novel fact for that page), and (b) to enable the researcher to convey, in a clear and succinct manner, the label of each picture (i.e., inclusion of distractor pictures would require labeling each picture individually, as well as subsequent checking to ensure that children remembered which pictures were labeled in which way).
The procedure for the clues-second condition was as follows. After the practice sets, the researcher opened the picture book, displaying the question and the animal picture set. The child was told which type of animal they were seeing (e.g., “Here is a page of bears”). The researcher then asked, “Which ones like to eat ants? Let’s look at the clue book to see if it can help us answer the question.” The researcher opened the clue book, read the corresponding clue, repeated it, and then asked the child to repeat it. The child was then told to peel each picture off and put it where it should go. They could place it either on the “yes” sheet if they believed the property was true of that particular animal or on the “no” sheet if they believed that the property was not true of that particular animal. This procedure was the same for all nine animal sets.

For the adult participants, the materials were compiled into a booklet, using the same orders and conditions. Participants were tested in small groups. They were informed that they were about to participate in a task that had also been conducted with very young children. Participants were also instructed to “assume, for the purposes of the task, that what we tell you is true, and answer on that basis.” After reading the instructions, the experimenter handed out a picture booklet and an answer sheet to each participant. The answer sheet provided a space for participants to circle “yes” or “no” for each test animal. Adults were allowed to work through the task at their own pace; all finished the task in less than 15 min.

The choice of noun quantifier was varied within subjects. Each participant received three clues in each of three conditions: all, generic, and some. The condi-
tions were blocked, and the order in which the conditions were presented was varied for each participant, with all six possible orders represented. All nine animals were divided into three groups of three animals. Order of presentation of the animal groups was fixed; within each group, the presentation sequence of the three animals was randomly varied. The animal groups were assigned to the conditions in all six possible orders.

Results
The number of "yes" responses was summed for each condition and typicality level; participants' scores ranged from 0 to 9. See Table 3 for results. We performed a four-way analysis of variance (ANOVA) on these scores: Age (adult, child) and order (clues-first, clues-second) were between-subjects variables, and wording (all, generic, some) and typicality (highly typical, moderately typical) were within-subjects variables.

There were three main effects. First, the wording main effect was significant, $F(2, 138) = 94.96, p < .0001$. Newman–Keuls post hoc analysis, $p < .01$, indicates that the induction rates for the all, generic, and some conditions were all significantly different from one another. Second, there was a main effect for age, $F(1, 69) = 11.31, p < .005$, indicating higher induction rates among the adults than among the children. Third, there was also a significant main effect of typicality, $F(1, 69) = 20.27, p < .0001$. The induction rate for highly typical category members was significantly higher than the induction rate for moderately typical category members (see Table 3).

Of greater interest were the interaction effects. There was a significant Age $\times$ Wording interaction, $F(2, 138) = 8.91, p < .0005$. Simple effects tests indicate that the age effect was localized on the generic items, with more inferences to generics by adults than by children, $F(1, 196) = 28.88, p < .0001$. Children and adults did not significantly differ on either the all or some conditions, $ps > .30$. Another way

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th></th>
<th>Generic</th>
<th></th>
<th>Some</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Children</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highly typical</td>
<td>8.32</td>
<td>1.31</td>
<td>7.19</td>
<td>1.73</td>
<td>5.73</td>
<td>2.04</td>
</tr>
<tr>
<td>Moderately typical</td>
<td>8.22</td>
<td>1.49</td>
<td>6.32</td>
<td>2.22</td>
<td>5.49</td>
<td>1.68</td>
</tr>
<tr>
<td>Adults</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highly typical</td>
<td>8.67</td>
<td>1.04</td>
<td>8.44</td>
<td>1.30</td>
<td>6.56</td>
<td>1.78</td>
</tr>
<tr>
<td>Moderately typical</td>
<td>8.44</td>
<td>1.36</td>
<td>8.42</td>
<td>1.25</td>
<td>4.94</td>
<td>1.97</td>
</tr>
</tbody>
</table>
of examining this interaction is to look separately at children and adults, with re-
spect to wording effects. Children drew significantly more inferences from “all”
than from generic, from generic than “some,” and from “all” than “some,” \( ps < .01, \)
as indicated by Newman–Keuls post hoc tests. In contrast, for adults, “all” and gen-
eric did not differ significantly from one another, and both were significantly
greater than “some,” \( ps < .01 \) (Newman–Keuls).

There was also a significant Wording \( \times \) Typicality interaction, \( F(2, 138) = 3.35, p < .05 \),
and an Age \( \times \) Wording \( \times \) Typicality interaction, \( F(2, 138) = 7.02, p < .005 \).
Simple effects analyses indicate that children displayed more inferences from
highly typical than moderately typical instances overall, \( F(1, 35) = 5.71, p < .05 \),
whereas for adults, typicality had a significant effect for “some” trials only, \( F(1, 34) = 15.05, p < .001 \).

Finally, there was a Wording \( \times \) Order interaction, \( F(2, 138) = 6.31, p < .005 \).
This interaction indicates that there were order effects for “all,” \( p < .05 \), and
generics, \( p < .01 \), but not for “some,” \( p > .15 \). As predicted, induction rates were
slightly higher for both “all” and generics when clues were provided second than
when clues were provided first. It appears that having the question in mind (as was
done in the clues-second condition) helps participants focus on the wording of the
clues.

None of the other interactions were statistically significant.

**Item Analyses**

To determine how consistent these results were over items, we looked at per-
formance on each of the nine-item sets considered separately. Children displayed
much consistency. On each of the picture sets, their mean “all” score was greater
than their mean “some” score; on eight out of nine sets, their mean “all” score was
greater than their mean generic score; and for eight out of nine sets, their mean ge-
neric score was greater than their mean “some” score. Adults also showed great
consistency in some respects. For adults, on all nine picture sets, both the mean ge-
neric score and the mean “all” score were greater than the mean “some” score.
However, the items were split with regard to the comparison of “all” and generic:
The mean “all” score was higher than the mean generic score on five of the nine
sets, and lower on four of the sets. Overall, then, the analysis of individual items
further supported the analyses shown previously with ANOVAs.

**Discussion**

Study 1 demonstrates that, for both preschool children and adults, inductive iner-
ences are sensitive to the language used when introducing category-relevant infor-
ination. The scope with which a novel property is expressed (i.e., with a generic
noun phrase, with a specifying noun phrase using “some,” or with a universally
quantified noun phrase using "all") affects the extent to which children and adults generalize that property to other category instances. It is noteworthy that 4-year-old children attended to this linguistic variation and used it to guide their inferences. Specifically, hearing a property attributed to "all" members of a category (e.g., "all bears") leads to more inferences than hearing a property attributed to a specific subset of a category (e.g., "some bears") for both children and adults. Of particular interest, use of a generic (e.g., "bears") was intermediate between (and distinct from) "all" and "some" for children, but equivalent to "all" for adults.

It is possible that children’s performance could have been affected by a failure to include one or more pictures in the animal category under consideration. For example, in trials using the category "bears," a child may have responded "no" because he or she felt that a particular picture was not a member of the category "bear," rather than because of the linguistic information provided by the clue. We do not believe that this is the case, however, as we designed the study to ensure that all pictured instances belonged to the category in question (e.g., that all six of the depicted bears were in fact bears). We took three precautions in this regard. First, we included only typical and moderately typical exemplars—and not atypical exemplars—so as to maximize the possibility that all items would be accepted as category members by preschoolers. Second, we included a pretest to find out which pictures children believed to be instances of the relevant categories. As detailed earlier, on 95.2% of the pretest trials the exemplars were selected as representing the category in question. Third, we labeled all instances before posing the test question (e.g., "Here is a page of bears"), to reinforce that all pictured instances were members of the category. Altogether, these three precautions render it unlikely that performance can be attributed to failure to include one or more instances to the category.

In addition to the cross-age similarities, there were important differences between children and adults in their interpretation of generic noun phrases. For adults, generics were treated as equivalent to "all," whereas for children, generics were treated as less powerful than "all" but more powerful than "some." These developmental differences suggest two conclusions: (a) Children are sensitive to generic noun phrases as fostering inductive inferences broadly within a category; and (b) preschool children favor a more conservative interpretation of generic noun phrases than do adults.

Item typicality also affected performance overall, with greater induction rates to highly typical items than moderately typical items. However, the importance of typicality varied by age. For adults, typicality affected performance only on the "some" trials. For children, typicality came through overall. Thus, evidence exists that generics are more likely to refer to highly typical category instances for children, but not for adults.

It was somewhat surprising that adults in Study 1 found generics to be as powerful as "all" statements in promoting category-based induction. This finding con-
trasts with analyses of generics in the linguistic literature, which suggest that semantically generics allow for exceptions, whereas “all” statements do not (Carlson & Pelletier, 1995; Dahl, 1975; McCawley, 1981; Prasada, 2000). Therefore, we sought additional data to help clarify adults’ interpretation. Specifically, we hypothesized that the task used in Study 1 may not have been sensitive enough to detect slight differences in quantifier scope, and that a more sensitive task would reveal differences between “all” and generics. Study 2 was designed to examine this possibility in adults. Because children already display a distinction between “all” and generics in Study 1, they were not included in Study 2.

STUDY 2

In Study 2, adults learned novel facts, as in Study 1, but then were asked to rate what percentage of the category had the property in question. For example, if respondents learned “Bears like to eat ants,” they were asked to indicate what percentage of bears like to eat ants. There are two features of this task that make it more sensitive than the task used in Study 1. First, participants in Study 2 were asked to consider what percentage of the entire category had the property in question. This contrasts with Study 1, in which all test pictures were either highly or moderately typical of the target category. In other words, only in Study 2 were atypical instances under consideration. It is with such atypical exemplars that generics should most strongly diverge from “all” statements. For example, the generic statement “Birds fly” applies uniformly to typical or moderately typical birds (and thus would be indistinguishable from “All birds fly” when considering such items), but does not uniformly apply to atypical birds (e.g., penguins, ostriches, and emus).

Second, the scale in Study 2 is more fine grained than the scale in Study 1, thereby allowing us to detect a smaller percentage difference. Specifically, participants in Study 1 responded on what was essentially a 7-point scale (selecting from 0 to 6 items as having the property in question). In contrast, participants in Study 2 responded on what was essentially an 11-point scale (ranging from 0% to 100%, in increments of 10%).

Of interest, therefore, is whether adults would display a distinction between “all” and generics when given a task designed to be more sensitive to small differences.

Method

Participants

Nineteen adults participated in this experiment. All were volunteers who received a gift certificate to a local ice-cream shop in exchange for participating.
Materials

The same categories and properties as used in Study 1 (see Table 1) were used in this study.

Procedure

Each participant received a questionnaire with 27 items. Each item was a statement in either all, generic, or some form, followed by a rating scale that went from 0% to 100%, in increments of 10% (i.e., 0%, 10%, 20%, 30%, etc.). For each item, the participant was asked to read the statement and judge what percentage of the category to which the statement applied.

Each participant saw the nine properties from Table 1 in all three forms (all, generic, and some), with wording blocked and in counterbalanced order. For example, some participants received nine “all” statements (e.g., “All bears like to eat ants”), followed by nine generic statements (e.g., “Bears like to eat ants”), followed by nine “some” statements (e.g., “Some bears like to eat ants”). All six possible orders were used.

Results and Discussion

Each participant received three scores, which were the averages of their ratings for “all,” generic, and “some.” These scores were entered into a two-way ANOVA, with order (“all” first, generic first, or “some” first) as a between-subjects variable, and wording (“all,” generic, or “some”) as a within-subjects variable. Results indicate a main effect of wording, $F(2, 32) = 76.17, p < .0001$, and no effects involving order. Newman–Keuls post hoc tests indicate that ratings for each type of wording differed significantly from each of the others: “all” (100%), generic (84%), and “some” (39%), $ps < .01$.

This measure indicates that adults do maintain a small but reliable semantic distinction between “all” and generics. Although “all” refers to all category members without exception, generics are interpreted as referring to the vast majority of the category.

Item Analyses

To determine how consistent these results were over items, we looked at performance on each of the nine item sets considered separately. Respondents were highly consistent. On all nine picture sets, the mean “all” score was higher than the mean generic score, and the mean generic score was higher than the mean some score. Specifically, the mean “all” score was 100%, for all nine items; the mean generic score ranged from 78.9% to 89.1%; the mean “some” score ranged from
33.6% to 42.7%. Overall the analysis of individual items further supported the analyses presented previously with ANOVAs.

**Individual Participants**

We also examined performance of individual participants, to determine inter-individual consistency. In broad strokes, the task elicited high agreement across individuals. Every participant interpreted “all” as referring to 100% of the category; 16 out of 19 individuals (84%) interpreted “some” as referring to less than one half of the category; and 17 out of 19 individuals (89%) interpreted generics as referring to more than one half of the category. However, when the data are examined more closely, individual participants showed interesting variation in their interpretation of generics. Although nearly one half of the participants (8 out of 19, or 42%) interpreted generics as referring to 100% of the category and thus indistinguishable from “all,” the remaining participants interpreted generics as allowing for exceptions, with mean scores ranging from 33.6% to 98%. These data imply that generics may be ambiguous, with some speakers interpreting them as equivalent to “all,” and others interpreting them as squarely distinct from “all.”

Because the individual differences analysis revealed nonnormality in the responses, we further conducted a set of nonparametric tests on the data. Wilcoxon signed-ranks tests reveal significant group differences between “all” versus generic ($T = 0, p < .01$), generic versus “some” ($T = 0, p < .01$), and “all” versus “some” ($T = 0, p < .01$).

**GENERAL DISCUSSION**

How do children learn to constrain or foster their inductive inferences? These studies demonstrate that for both preschool children and adults, language powerfully influences the scope of an inference. When novel information is predicated of “some” category instances, relatively few inferences are made, whereas when novel information is predicated of “all” category instances or the generic kind, many more inferences are made. It would not have been surprising if 4-year-olds had ignored the linguistic cues provided and focused simply on picture typicality. We found instead that both children and adults consistently distinguished these various language forms.

The variation in form and meaning across the conditions of these studies was subtle in two respects. First, the distinction between “all” and generics is subtle in that both noun phrase types imply the category as a whole, differing only in their commitment to every instance. At times generics are indistinguishable from “all,” being used to convey properties true of every category member (e.g., “Dinosaurs are extinct”). Nonetheless, generics allow for the possibility of counterexamples in a way that “all” does not. Second, the distinction between generics and “some” is
subtle in that the bare plural form has dual functions, at times referring to generics (e.g., "Bears like to eat ants") and at times referring to an indefinite plural (e.g., "I saw bears in the park yesterday"). Note that the indefinite plural is comparable to a noun phrase with "some" (i.e., "I saw bears in the park yesterday" is comparable in meaning to "I saw some bears in the park yesterday"). Given these potential confusions—on the one hand between generics and "all," and on the other hand between generics and "some"—it is all the more impressive that 4-year-olds consistently distinguished among the three forms.

Generics are of particular interest because of their frequency in child-directed speech, their semantic implications (as being generally true of a category and yet allowing for exceptions), and—until now—the lack of data concerning how they are understood by children. The results from both studies demonstrate that generics are not simply interpreted as equivalent to "some" statements (despite their allowance of exceptions), nor are they equivalent to "all" statements (despite their generality). Rather, for children they seem to be intermediate between "some" and "all."

This result is relevant to earlier findings regarding the use of generics in mother–child interactions, as it tells us that children are capable of using such information to drive or shape their conceptual representations. In other words, there is a dovetailing of parental speech and child expectations.

In addition to revealing that generics elicit consistent category-based inductive inferences, the data indicate an age difference in the generics condition, when comparing preschool children with adults. Specifically, 4-year-olds display a more conservative interpretation of generics, showing less willingness to generalize from a generic statement to the category as a whole. In contrast, adults treat generics as nearly equivalent to "all" statements. We do not yet know the reason for the age difference. There may be change with age in the semantics of generics, or it may be that children and adults are differentially sensitive to contextual factors that highlight the "all" and generics distinction. For example, the studies presented here explicitly blocked the items; thus potentially exaggerating the distinction between the different linguistic forms. It is possible that the young children were more sensitive to the pragmatic implications of the resulting contrast.

It is also unclear at what age performance on this task changes from the more conservative reading favored by 4-year-olds to the more generous reading favored by adults. However, we speculate that the caution displayed by the children may be adaptive, given the frequency with which generics appear in parental speech. Children's awareness that generics admit exceptions could help prevent overly broad generalizations.

Another issue that remains unresolved concerns the interpretation of generic noun phrases across content domains. In these studies, all information concerned familiar animal categories. This choice was deliberate and reflects the fact that generics in natural speech are more commonly used to refer to people or animal kinds than to artifact categories (Gelman et al., 1998). In the future, it would be in-
teresting to test whether generic interpretation likewise differs across domains. For example, perhaps generics elicit broader generalizations when applied to animals (e.g., “Bears like to eat ants”) than when applied to artifacts (e.g., “Chairs are eaten by ants”). This would follow from the fact that natural kinds tend to possess rich, correlated structure that artifacts lack (Keil, 1989; although “all” statements would also seem more felicitous for animal properties than artifact properties, we suspect that the explicitness of “all” would encourage a fairly comparable interpretation across domains; this, however, remains an open empirical question).

A cross-domain comparison would present some methodological difficulties, however, because intuition suggests that it would be difficult to construct a set of novel properties for artifacts that sensibly fit a generic sentence frame. Artifact properties that readily fit the generic sentence frame tend to be highly familiar (e.g., “Chairs are used for sitting; knives are sharp”). On the other hand, unfamiliar properties placed in the generic sentence frame can seem implausible. For example, “Chairs are made out of substance X” (whatever X may be) may be rejected as infelicitous because people realize that chairs vary considerably in substance and materials.

We turn to some speculations concerning the role of generics in knowledge construction. There are two distinct but related purposes that generic statements may serve (Gelman et al., 1998; Pappas & Gelman, 1998). First and most obviously, generics are effective in teaching children particular category-wide generalizations, such as specific facts about the attributes of category members. The data from these studies suggest that, by means of generic noun phrases, children learn particular facts about the physical characteristics, eating habits, and behaviors of animals. Because such facts are often stated generically, they may “become more central to children’s conceptual representations than if they had been stated non-generically” (Gelman et al., 1998, p. 84). This was the focus of these studies, and the results support this interpretation.

Second, the use of generics in maternal speech may indicate to children that a category as a whole is “an inference-promoting entity” (Gelman & Diesendruck, 1999, p. 84). Through the use of generics, categories may become richly structured and amenable to inductive inferences. In other words, generics may aid in the development of categories as kinds (Carlson & Pelletier, 1995). The suggestion here is that referring to a category in generic form implies that the category is information-rich and inductively powerful. When children hear a statement such as, “Birds lay eggs,” they learn not only a particular fact (about egg laying), but also an implication about birds as a category—that “birds” cohere into a category about which general, law-like regularities accrue. The generic statement (i.e., “Birds lay eggs”) “presupposes the conceptualization of the class of birds as a single entity...[and therefore] should enhance the psychological coherence of the class of birds for that reason” (Shipley, 1993, p. 278). Indeed, if this line of reasoning is correct, then even relatively content-free generic statements (e.g., “I like birds; What do you
think of birds?"") could imply the richness of birds as a kind. These speculations await future research.

Finally, we point out that, although language was the focus of these studies, it is not the only source of information that either children or adults use to guide their inferences. As Callanan noted (1990), "Children’s inferential abilities are likely to develop from a complex interaction among their theories and expectations about the properties and categories involved, the knowledge they gain through their senses, and the knowledge they gain through verbal descriptions (most notably provided by parents)" (p. 106). In future work, it will be of interest to explore these interactions among different sources of knowledge.

ACKNOWLEDGMENTS

This research was supported by NICHD Grant HD36043 awarded to Susan A. Gelman. We are grateful to the teachers, staff, and children of the University of Michigan Children’s Center, University of Michigan Center for Working Families, Peach Tree Day Care, Little Tiger Day Care, Jewish Community Center of Ann Arbor, Perry Nursery School, Stoney Creek Preschool Too, and University Child Development School of Seattle for participating in this research. We thank Andrew Meltzoff for generously helping us contact preschools in Seattle during pilot testing and Matthew Germak for his able assistance in testing.

REFERENCES


