Published in final edited form as:
Dev Sci. 2010 May ; 13(3): 421–429. doi:10.1111/j.1467-7687.2009.00896.x.

Children's attention to sample composition in learning, teaching, and discovery

Marjorie Rhodes, **Susan A. Gelman**, and **Daniel Brickman**
University of Michigan

Abstract

Two studies compared children's attention to sample composition—whether a sample provides a diverse representation of a category of interest—during teacher-led and learner-driven learning contexts. In Study 1 ($N = 48$), 5-year-olds attended to sample composition to make inferences about biological properties only when samples were presented by a knowledgeable teacher. In contrast, adults attended to sample composition in both teacher-led and learner-driven contexts. In Study 2 ($N = 51$), 6-year-olds chose to create diverse samples to teach information about biological kinds to another child, but not to discover new information for themselves, whereas adults chose to create diverse samples for both teaching and information discovery. Results suggest that how children approach the interpretation and selection of evidence varies depending on whether learning occurs in pedagogical or non-pedagogical contexts.

Children's attention to sample composition in learning, teaching, and discovery

Much of human learning involves extending information obtained from limited samples to inform general expectations about the world. Some of this learning occurs in pedagogical contexts, in which one person, the *teacher*, knows some target information that another person, the *learner*, does not (Csibra & Gergely, 2006). Pedagogical learning does not require a formal teacher, but instead is defined by the epistemic gap between the teacher and learner and the intent of the teacher to communicate information to the learner. In contrast, other learning occurs in the absence of someone who possesses the target knowledge (referred to here as non-pedagogical learning). The goal of such learning is information *discovery*, and involves processes that are learner-driven, including observation and induction (e.g., after observing one bird build a nest, learners may infer that, generally, birds build nests; see Feeney & Heit, 2007), as well as evidence gathering, in which learners seek evidence to evaluate whether inferences are accurate (e.g., Rhodes, Gelman, & Brickman, 2008).

The goal of the present studies was to examine the role of sample composition—particularly, whether a sample provides a diverse representation of a category of interest—in children's use of evidence in pedagogical and non-pedagogical learning. Following much prior work on conceptual development, we focused on children's learning about biological kinds. To preview our hypotheses, we propose that an appreciation of sample composition emerges earlier in pedagogical than in non-pedagogical learning. Such developmental differentiation would suggest that the learning context importantly influences how children

Address for correspondence: Marjorie Rhodes, Department of Psychology, University of Michigan, 2422 East Hall, 530 Church Street, Ann Arbor, MI 48109-1043. rhodesma@umich.edu; fax: (734) 615-0573.

make sense of evidence, and would help to clarify the strengths and limitations of children's reasoning about sample properties.

The Role of Sample Composition in Pedagogical Interactions

Efficient teachers purposefully select evidence to create samples that clearly and unambiguously represent concepts of interest (referred to as *pedagogical sampling*, Shafto & Goodman, 2008). For example, if a teacher wants to teach about a property of birds, it seems more effective to present a sample containing three different birds (e.g., a canary, a peacock, and an eagle), than a sample containing only one kind of bird (e.g., three canaries). The latter sample is ambiguous regarding whether the property applies to all birds or only to canaries, whereas the diverse sample more efficiently communicates that the property applies to all birds. Shafto and Goodman (2008) present a computational model and empirical evidence demonstrating that adults assigned to teaching roles readily engage in such pedagogical sampling, without explicit instruction to do so.

Prior work also suggests that learners *assume* that teachers engage in helpful sampling methods, and that this assumption facilitates learning. For example, Xu and Tanenbaum (2007a) found that when a knowledgeable puppet presented preschoolers with an array of different kinds of dogs and selected three of the same type (e.g., three dalmatians) to label as "blickets", children generalized the term "blicket" only to the subordinate class. Alternately, when the puppet labeled one exemplar each of three different kinds of dogs as "blickets", children extended the term to the basic level. Ellis, Denison, and Xu (2007) report similar findings regarding property induction; four-year-olds generalized new properties to the subordinate level after being taught on a sample from one subordinate class, but to the basic level after being taught on a sample containing three different subordinates. As described by Xu and Tanenbaum (2007b), the assumption that samples were selected *purposefully* was critical to children's learning. Indeed, in a word learning study similar to those described above, when preschoolers believed that samples were created by someone who did not *know* the true meaning of "blicket", they no longer used sample composition to determine the extension of the word. Thus, children's interpretation of evidence depended on whether they believed samples were formed with specific knowledge and intent.

During pedagogical learning, sample composition provides a window into the communicative goal of the teacher. According to Csibra and Gergely (2006), reasoning about communicative goals may be facilitated by intuitive biases that support pedagogical exchanges, which evolved to support the rapid teaching of information (see Gergely, Egyed, and Király, 2007). The extension to interpreting sample composition is straightforward; from a child's perspective, for example, "The teacher is trying to teach me something. She had all these dogs to choose from and chose three different kinds (or three of the same kind). That decision was purposeful, so she must be trying to teach about all dogs (or just one kind of dog)". Thus, children's early emerging abilities to reason about sample composition in pedagogical contexts may stem from an intuitive sense of how sampling behavior reflects communicative goals (see Shafto & Goodman, 2008).

The Role of Sample Composition in Non-Pedagogical Learning

In non-pedagogical learning, learning goals involve information *discovery*, instead of information *communication*. In these learner-driven processes, diverse samples provide stronger tests for hypotheses about the world (Heit, Hahn, & Feeney, 2004). As described by Osherson and colleagues (1990), for example, if a person wants to discover whether birds all have a particular enzyme, testing whether a penguin and a sparrow both have the enzyme would provide stronger evidence than would testing two sparrows. The first sample presents stronger evidence because sparrows and penguins are different from each other in many

NIH-PA Author Manuscript

NIH-PA Author Manuscript

other ways, whereas two sparrows share many properties that are specific to their subordinate class (Heit, 2000). Prior empirical work indicates that adults appreciate the value of sample diversity for information discovery; they choose to create diverse samples when asked to test hypotheses about categories (Lopez, 1995; Rhodes, Brickman, & Gelman, 2008; Rhodes, Gelman, & Brickman, 2008), and they are more willing to generalize information beyond observed samples when samples include diverse, as compared to non-diverse, exemplars (e.g., Gutheil & Gelman, 1997; Lopez, Gelman, Gutheil, & Smith, 1992; though adults consider factors beyond diversity when evaluating samples as well, see Medin, Coley, Storms, & Hayes, 2003).

Prior work with children suggests, however, that consideration of sample composition in non-pedagogical learning is a relatively late development (Carey, 1985; Gutheil & Gelman, 1997; Lopez, et al., 1992; Rhodes, Brickman, & Gelman, 2008; Rhodes, Gelman, & Brickman, 2008; for competing perspectives see Heit & Hahn, 2001; Shipley & Shepperson, 2006). For example, Gutheil and Gelman (1997) found that nine-year-olds were equally likely to generalize a novel property to a category (e.g., birds) based on exemplars from a single subordinate (e.g., 5 robins) as on exemplars from different subordinates (e.g., 5 different kinds of birds; no pedagogical cues were presented in this study). Tasks examining children’s evidence selection have yielded similar results; for example, when asked to determine if there is support for a generalization (e.g., about all dogs), six-year-olds were equally likely to choose to examine two animals of the same type (e.g., two dalmatians) as two animals of different types (e.g., a dalmatian and a golden retriever; Rhodes, Brickman, & Gelman, 2008). Thus, in contrast to the findings on pedagogical learning, children appear not to attend to sample composition in non-pedagogical learning until relatively late in childhood.

We propose that the apparent differences in children’s use of sample diversity for pedagogical and non-pedagogical learning relates to conceptual differences in the significance of sample composition across different types of learning. As reviewed above, pedagogical learning involves the transmission of established information; in such contexts, sample diversity provides a cue to a teacher’s communicative goal, and use of sample composition as a window into communicative goals may be supported by intuitive biases that support pedagogical exchanges. In contrast, non-pedagogical learning involves the discovery of new information; in these contexts, an appreciation of sample diversity rests on the belief that some samples provide stronger tests than others for hypotheses about the world (Heit, 2000). This belief may be a developmental achievement, instead of an intuitive bias. Young children often overlook variability within categories (Gelman, 1988) and instead treat individual category members as standing in for a category as whole (Gelman, 2003). If indeed early categories do not include a focus on variability within a kind (e.g., if early concepts focus on the similarities that birds share, as opposed to the variability among birds), children may not recognize the value of obtaining diverse representation from a category before making category-wide generalizations.

In sum, although incorporating sample composition into learning requires many of the same cognitive skills in pedagogical and non-pedagogical contexts (e.g., similar information-processing skills), prior work suggests that the meaning and implication of sample composition differ across the two types of learning contexts. The goal of Study 1 was to test the hypothesis that young children attend to sample composition to guide their learning in pedagogical, but not non-pedagogical, contexts. As described above, prior empirical work supports this hypothesis. Also as reviewed above, there are theoretical reasons to expect children to attend to sample composition in pedagogical exchanges (e.g., when sample composition provides a cue to a communicator’s intent) at a younger age than they attend to sample composition in non-pedagogical learning (e.g., when sample composition relates to how well samples test hypotheses about the world). However, because children's reasoning about sample properties in different learning contexts has been previously examined in separate studies using substantially different methods, the goal of Study 1 was to provide a direct comparison.

Study 1

In a between-subjects design, we manipulated whether teaching samples were presented by someone who knew a lot about animals (expert condition), or a sampler who knew nothing about animals (novice condition; see Kushnir, Wellman, & Gelman, 2008). In the expert condition, learning occurred in a pedagogical context; participants believed that samples were intentionally created by a knowledgeable sampler with a communicative goal. In the novice condition, learning occurred in a non-pedagogical context; the sampler "discovered" the information while presenting the sample. Thus, the novice provided *accurate* information, but was not aware of the true property distribution ahead of time, and therefore could not have purposefully selected the sample to facilitate a communicative goal. In both conditions, the sampler sometimes presented non-diverse teaching samples (e.g., three dalmations), and sometimes presented diverse teaching samples (e.g., a dalmatian, a collie, and a basset hound). We tested whether participants generalized properties that were described in the teaching sample to other exemplars that matched at the subordinate level (e.g., another dalmation), as well as to exemplars that matched at the basic level, but were not included in the teaching set (e.g., a golden retriever).

For generalizations to the subordinate level, there should be no effects of sample (diverse vs. non-diverse) or condition (expert vs. novice) on children's or adults' property extensions; they should reliably extend properties to these subordinate matches. For generalizations to the basic level, however, for adults, there should be an effect of teaching sample in both conditions. Particularly, in the expert condition, the diverse sample indicates that the teacher intends to teach about the basic level (whereas non-diverse samples indicate the subordinate level), and in the novice condition, the diverse samples provide strong evidence for generalizations (whereas non-diverse samples provide weak evidence). Therefore, in both cases, diverse samples should lead to increased property extensions to basic-level matches. In contrast, we predicted a condition X sample interaction for children, such that they will show an adult-like pattern in the expert condition (where diversity relates to communicative intent), but not in the novice condition (where diversity relates to evidential strength). We selected kindergarteners because 5-year-olds are well within the age range of children who have been shown to attend to sample composition in pedagogical contexts (Xu, 2007), but are younger than children who have been shown to incorporate sample composition into induction involving biological kinds (e.g., Gutheil & Gelman, 1997).

Method

Participants—Children (N = 28; M age = 5.3 years, range = 4.9–5.9; 17 male, 11 female) were recruited from kindergarten classrooms in a public elementary school. Only children who returned signed parent consent letters were asked to participate. College students (N = 20; 7 male, 13 female) were undergraduate students who volunteered to participate in exchange for a $5 gift card. Equal numbers of participants of each age were randomly assigned to the expert or novice condition.

Procedures—Participants completed the task individually with a trained female research assistant, children in a quiet area at their school, and adults in campus libraries. All participants completed two trials: a diverse sample trial and a non-diverse sample trial. The visual display consisted of 7 rows of animal photographs; each row contained three exemplars of one type of animal (see Table 1). The first five rows consisted of five types of dogs (e.g., dalmatians, golden retrievers, black labradors, collies, basset hounds). The sixth row consisted of three monkeys, and the seventh row consisted of three turtles. The following factors were controlled by counter-balancing across participants: whether the diverse or non-diverse sample trial was presented first, which exemplars from the category DOG were teaching exemplars and which were test items (and the order of the dog categories in the visual display), and which novel property (epithelium, four-chamber heart) was described during the diverse or non-diverse sample trials.

Expert Condition: In the expert condition, children were introduced to a human-like puppet and told, "This is Daxy. Daxy knows a lot about animals. He wants to tell you some new things that he knows, and then he'll ask you to make some guesses." For the teaching phase, the experimenter first explained the properties of interest. For the property *epithelium*, the experimenter said, "First Daxy wants to teach you about something called an epithelium. Can you say epithelium? Great! Now, let me tell you what it is. Do you know what lungs are? Lungs are the things inside of us that help us breathe [experimenter took a deep breath to show lungs working]. Well, animals have lungs too! Some animals have something called an epithelium inside their lungs that helps them work. Some animals have it and some animals don't. Daxy is going to teach you which animals have an epithelium. He'll show you three animals, and after that, you'll have to guess." A parallel property explanation was provided for the property "four-chamber heart".

After the property explanation, the experimenter pointed to the first teaching exemplar and said, "Daxy says, 'See this one? It has an [epithelium/ four-chamber heart]!'", and repeated this sentence for each of the two remaining teaching exemplars. On the non-diverse sample trial, the teaching exemplars were the first three animals in the top row (e.g., three dalmations). For the diverse sample trial, the teaching exemplars were the first animal in each of the first three rows (e.g., a dalmation, a golden retriever, and a black labrador). To begin the test phase, the experimenter said, "Now Daxy says it is your turn to make some guesses." The experimenter presented a series of six test items, in the following set order (following Xu & Tanenbaum, 2007b): a subordinate-level match (e.g., a dalmatian), an animal from another basic-level category (e.g., a monkey), a basic-level match from a subordinate level not included in teaching (e.g., a collie), another subordinate-level match (e.g., a dalmatian), another basic-level match from a subordinate level not included in teaching (e.g., a basset hound), and an animal from a more distant category (e.g., a turtle). For each test item, the experimenter asked, "Does this one have an [epithelium/four-chamber heart]?" Children responded by saying "yes" or "no".

After the first trial, the experimenter said, "Now Daxy wants to tell you some more things", and provided the property explanation for the second novel property. Then, the experimenter completed the teaching phase using the same verbal script as in the first trial, but with the alternate teaching sample (e.g., the diverse teaching sample if the first trial included the non-diverse sample) and the other novel property.

For adults, the sampler was described as "an expert who knows a lot of things about animals" and was referred to as "the expert" throughout the study. The wording of the property explanations was also revised to be age appropriate.

Novice Condition: In the novice condition, the same puppet was used for children, but given a different introduction: "This is Daxy. Daxy doesn't know anything about animals. He wants to learn some new things. He'll learn some new things by looking inside the animals with his special eyes. Then, he'll ask you to make some guesses." The properties were introduced with the same explanations as in the expert condition. Then, the experimenter pretended to have the puppet look inside each of the teaching exemplars, while saying “Daxy looks in this one and says, ‘Oh boy! It has [an epithelium/ a four-chamber heart] inside!’” The rest of the procedure was identical to the expert condition. For adults, the sampler was described as “a visitor from some place far away who doesn’t know anything about animals” and was referred to as “the visitor” throughout the study.

Results

Children—We compared the proportion of property extensions to the proportion expected by chance, separately by condition and teaching sample. Then, to examine effects of condition and teaching sample on property extensions, we conducted 2 (sample: diverse, non-diverse) X 2 (condition: novice, expert) factorial analyses, with sample as a within-subjects factor, separately for subordinate and basic-level extensions. All analyses were conducted using the generalized estimating equations (GEE) procedure in SPSS 16, which is appropriate because the underlying structure of the data was binary – the dependent variable was analyzed as the number of “yes” responses out of the total number of questions (2 for subordinate level matches and 2 for basic level matches). Also, this procedure can account for both between- and within-subjects variables. These analyses yield Wald χ^2 values as indicators of the significance of main effects and interactions. Descriptive statistics are reported as proportions of “yes” responses.

The Subordinate Level: Children reliably extended the novel properties to animals that matched the teaching exemplars at the subordinate level (e.g., two dalmations) in both conditions, following both non-diverse (Expert, *M* = .86, *SE* = .10; Novice, *M* = .89, *SE* = .08) and diverse, (Expert, *M* = .86, *SE* = .08; Novice, *M* = .93, *SE* = .05) teaching trials, *p*s < .05. For generalizations to the subordinate level, there were no effects of condition or teaching sample (*p*s > .4).

The Basic Level: In the expert condition, children reliably extended the novel properties to animals that matched the teaching exemplars at the basic-level (but were not included in the teaching sample) only when they were taught on diverse teaching samples (*M* = .86, *SE* = .10, $p < .05$). When children in the expert condition were taught on non-diverse teaching samples, they reliably failed to generalize the property to basic-level matches (*M* = .21, *SE* = .10, $p < .05$). In contrast, in the novice condition, children extended to the basic-level marginally more often than expected by chance, $p < .07$, (Non-diverse, *M* = .79, *SE* = .11; Diverse, *M* = .68, *SE* = .11).

For generalizations to the basic level, there was an interaction between condition and sample, $\chi^2(1) = 15.35$, $p < .001$, see Figure 1. Post-hoc contrasts, with sequential Bonferroni corrections, revealed that in the expert condition, children generalized to basic-level matches significantly more often following diverse, as compared to non-diverse, teaching samples, $p < .01$, whereas in the novice condition, children’s responses did not differ by teaching sample, $p > .4$. Also, the effect of condition was specific to responses following non-diverse teaching trials; for these trials, children in the expert condition generalized significantly less often than children in the novice condition, $p < .01$. In contrast, responses following diverse sample trials did not differ by condition, $p > .4$.

The Superordinate Level: Overall, children did not reliably extend the properties to animals that were more distantly related to the teaching exemplars (Monkeys, *M* = .34, *SE* = .07; Turtles, *M* = .30, *SE* = .07). The only effect of condition was for extensions to monkeys following non-diverse teaching trials; on these items, children extended less often in the expert condition than in the novice condition ($p < .05$, Fisher’s exact test).

Analyses of Individual Response Patterns—Participants were coded as "subordinate-only" if they extended the property to both subordinate-level matches and no basic-level matches; "basic-level" if they extended to both subordinate matches and both basic matches; and "other" if they had an alternate pattern of responding (e.g., extended to one subordinate and one basic). These codes are presented in Table 2. In the expert condition, the distribution of these codes differed by sample, Fisher's exact test, $p < .01$. The most common code following non-diverse samples was subordinate-only, whereas the most common code following diverse samples was basic-level. In the novice condition, the distribution of individual response patterns did not differ by sample, $p > .9$. The most common code following both diverse and non-diverse samples was basic-level. We also calculated the number of children who displayed the normatively predicted pattern of subordinate-only following non-diverse teaching samples and basic-level following diverse samples. In the expert condition, 8 children fit these criteria, whereas in the novice condition, 0 children met these criteria ($p < .01$, Fisher's exact test).

Adults—In several cases there was no variability in adults' responses; therefore, the GEE procedure could not be used to assess effects of condition and sample. Instead, we used non-parametric Mann-Whitney tests for independent samples to assess condition effects and Wilcoxon signed-rank tests for related samples to evaluate the effects of sample.

The Subordinate Level: Adults reliably extended to subordinate-level matches (Expert condition: Non-diverse, $M = .95$, $SE = .05$, Diverse, $M = 1$, $SE = 0$; Novice condition: Non-diverse, $M = 1$, $SE = 0$, Diverse, $M = 1$, $SE = 0$), *p*s < .01. There were no effects of condition or sample, *p*s > .3.

The Basic Level: Adults reliably extended to the basic level following diverse teaching samples in both the Expert ($M = 1$, $SE = 0$) and Novice ($M = .90$, $SE = .07$) conditions, *p*s < .01. However, following non-diverse teaching trials, they extended to the basic level less often than expected by chance, $p < .05$ (Expert, $M = .30$, $SE = .13$; Novice, $M = .30$, $SE = .13$). There were no effects of condition, *p*s > .14, see Figure 1. There was an effect of sample within both conditions, such that adults extended to the basic level more often following diverse than non-diverse teaching trials in both the expert, $Z = 2.64$, $p < .01$, and novice, $Z = 2.59$, $p = .01$, conditions.

The Superordinate Level: Adults generally did not extend the properties to either monkeys ($M = .15$, $SE = .06$) or turtles ($M = .10$, $SE = .06$). There were no effects of condition, *p*s > .2.

Individual Response Patterns: Individual response patterns are presented in Table 1. Adults' individual response patterns differed by sample in both conditions, *p*s < .01.

Age Comparisons—To compare the performance of children and adults, we conducted a series of Mann-Whitney tests for extensions to the subordinate and basic level, following diverse and non-diverse teaching trials, separately by condition. In the expert condition, there were no differences by age, *p*s > .1. In the novice condition, the only age effect was for extensions to the basic level following non-diverse teaching trials. On these items, children were more likely than adults to extend to the basic level, $Z = 2.51$, $p < .05$.

Discussion

Young children attended to sample composition to guide their inferences when samples were presented by a knowledgeable teacher, but not during non-pedagogical learning. Their performance in the expert condition suggests that, for property induction, young children

NIH-PA Author Manuscript

NIH-PA Author Manuscript

attend to sample composition as a cue to communicative intent. Particularly, when a teacher presented a sample that represented a single subordinate class (i.e., the non-diverse sets), children interpreted the teaching experience as indicating that the property applies only to the subordinate level, whereas when a teacher presented a sample that diversely represented a basic level category, they extended the property to the basic level. In contrast, children's performance in the novice condition indicates that they did not recognize the value of sample diversity for evaluating evidence strength. Particularly, when samples were presented by unknowledgeable samplers, children extended to the basic level following both diverse and non-diverse teaching samples, suggesting that they did not view non-diverse samples as providing weaker evidence than diverse samples. Thus, 5-year-olds appear to recognize the role of sample composition in information communication, but not for information discovery.

Adults extended to the basic level only after diverse teaching samples, regardless of whether the sampler was described as knowledgeable. It is important to note that these findings for adults are specific to *property induction*, a process for which adults have general beliefs about the types of samples that are inductively rich. For example, we propose that adults' valuing of sample diversity in the non-pedagogical context rested on their belief that diverse samples provide stronger evidence for testing hypotheses about categories (Heit, 2000). In other areas, such as word learning, adults may not view sample diversity as informative beyond as a cue to communicative intent, and in these cases, they should be sensitive to sample composition only in pedagogical contexts (Xu & Tanenbaum, 2007b; Shafto & Goodman, 2008).

Study 2

In Study 2, we used an evidence-selection method to test the hypothesis that young children attend to sample properties in pedagogical but not non-pedagogical contexts. In one condition, participants were placed in the role of "teacher" and asked to select a sample to use to teach another child a fact about an animal category. In another condition, participants were placed in the role of "scientist", and asked to select a sample in order to find out something new about an animal category (e.g., Rhodes, Brickman, & Gelman, 2008). In both conditions, participants were asked to select between a diverse and a non-diverse set of evidence.

From a normative perspective, the diverse sets should be viewed as more valuable than the non-diverse sets in both conditions, but for different reasons. In the teacher condition, the value of sample diversity relates to information *communication*. Particularly, participants' goal is to communicate that a property is true for a basic level category (e.g., for all dogs). Because we expect both children and adults to recognize the value of sample composition for communication, we hypothesize that, across ages, participants will judge that samples that provide diverse representation of basic level categories will communicate the intended information more efficiently than samples drawn from a single subordinate class (as these non-diverse sets leave open the possibility that the property is true only at the subordinate level). In contrast, in the scientist condition, the value of sample diversity relates to information *discovery*. In this condition, we hypothesized that adults would reliably select diverse samples, as they value diverse samples as providing stronger tests for hypotheses about categories, but that children will fail to value sample diversity in this context because they do not recognize the value of obtaining diverse samples from categories before making category-wide generalizations.

Method

Participants—Participants included 31 kindergarteners (*M* age = 6.01 years, range = 5.25 – 6.58; 16 male, 15 female) recruited from a public elementary school, and 20 college students who volunteered to participate (6 male, 14 female). Only children who returned signed parent consent letters were asked to participate; no child had participated in Study 1. Participants of each age were randomly assigned to a Teacher (17 children; 10 adults) or Scientist (*n* = 14 children, 10 adults) condition.

Procedures—Participants completed two questions, one involving dogs, and the other involving birds. Children completed the task individually with a trained experimenter in a quiet area of their school; adults completed a paper-and-pencil version of the task independently. For each question, participants were shown two sets of animals, containing three exemplars each. Exemplars were colorful photographs of individual animals. One set contained exemplars from one subordinate class (e.g., three dalmations); the other set contained exemplars drawn from different subordinate classes (e.g., a golden retriever, a dalmation, and a collie).

Teacher: In the teacher condition, on the dog trial, children were told, “Look at all the dogs! Let’s pretend that you are a teacher who is trying to teach someone about dogs. Your job is to teach this child [*show picture of a young girl*] that dogs have [four-chamber hearts]. But, you can’t show all the dogs in the world to teach that they have [four-chamber hearts], you can only show her three dogs. Just three. Your job is to pick out the three best dogs to show her to help her learn that dogs have [four-chamber hearts]. Which three dogs should you show her to teach about dogs?” Half of the children were asked to teach that dogs “have an epithelium inside” instead of “four-chamber hearts.”

Scientist: In the Scientist condition, for the dog trial, children were told, “Look at all the dogs! Let’s pretend that you are a scientist who is trying to learn about dogs. Your job is to find out if dogs have [four-chamber hearts]. But, you can’t look at all the dogs in the world to find out if they have [four-chamber hearts], you can only look at three dogs. Just three. Your job is to pick out the three best dogs to look at to help you find out if dogs have [four-chamber hearts]. Which dogs should you look at to find out about dogs?”

In both conditions, instructions on the bird trial were identical to those described above (but referred to different novel properties: whether birds have gizzards inside, or have hollow bones). The following factors were controlled by counterbalancing across participants: whether the bird trial or dog trial was presented first, whether the diverse or non-diverse sample was presented on a participant’s left side, and which subordinates were included in the non-diverse sets. The diverse set was composed of one animal from each subordinate class (e.g., a collie, a golden retriever, a black lab). For birds, exemplars included cardinals, blue jays, and canaries.

Results

Children selected the diverse samples significantly more often than expected by chance in the teacher condition ($M = .74$, $SE = .08$, $p < .01$), but at levels predicted by chance in the scientist condition ($M = .50$, $SE = .07$). In contrast, adults selected diverse samples significantly more often than expected by chance in both conditions (M scientist = .85, $SE = .08$; M teacher = .90, $SE = .10$, ps < .05). Mann-Whitney tests revealed an effect of condition for children ($Z = -2.14$, $p < .05$), such that children were less likely to select the diverse sample in the scientist than in the teacher condition, but no effect of condition for adults ($p > .30$).

In the teacher condition, 9 children selected the diverse sample on both questions (7 selected the diverse sample on 1 question, and 1 child selected the non-diverse sample on both questions). In the scientist condition, 2 children selected the diverse sample on both questions (10 selected the diverse sample on one question and 2 children selected the non-diverse sample on both questions). The number of children who selected diverse samples on both trials differed marginally by condition, $p < .06$, Fisher's exact test. For adults, the number of people who selected the diverse sample on both questions did not differ by condition (9 in the teacher condition and 7 in the scientist condition).

Discussion

The findings from Study 2 suggest that kindergarteners understand the role of sample composition in teaching, and provide the first evidence that we are aware of indicating that young children systematically select useful samples when placed in teaching roles. Consistent with prior work, however, kindergarteners appear not to recognize the value of diverse evidence for information discovery. In contrast, adults reliably chose to select diverse samples both when they were asked to communicate information and when they were asked to discover information, suggesting that they value sample composition in both pedagogical and non-pedagogical learning.

General Discussion

In two studies, young children attended to sample composition in pedagogical, but not non-pedagogical learning. In Study 1, 5-year-olds attended to sample composition to determine how to generalize new knowledge about biological kinds, but only when samples were created by a knowledgeable teacher. In Study 2, kindergarteners attended to sample composition when asked to teach another child a new fact about a biological kind, but not when they were asked to discover new information about a biological kind for themselves. In both studies, children demonstrated markedly different performance depending on whether they were placed in pedagogical or non-pedagogical contexts, even though the tasks were highly similar across conditions. Together, these findings support the hypothesis that children's initial consideration of sample composition is embedded in pedagogical contexts, and that the importance of sample composition for induction and discovery emerges later in development.

These findings are consistent with recent theorizing that pedagogical contexts cue specific kinds of learning (Gergely et al., 2007; Shafto & Goodman, 2008). By 5 years, and even younger (see Xu & Garcia, 2008), children assume that teachers engage in purposeful sampling in pedagogical contexts, and use this assumption to facilitate learning. Engaging in pedagogical sampling when placed in teaching roles also appears to be a relatively early emerging skill. Thus, these findings are consistent with Csibra and Gergely (2006), who proposed that humans are natural teachers and learners, equipped with early emerging cognitive biases that support pedagogical exchanges.

In future work, it will be important to examine children's sampling assumptions more specifically. Xu (2007) suggested that children assume that teachers present *random* samples of a true concept. In many circumstances, a random sample will also yield the sample that most efficiently communicates the concept (e.g., if a teacher randomly samples three dogs from 5 different kinds of dogs, or purposefully selects three different kinds of dogs from the same set, the results will usually be similar). Shafto and Goodman (2008) point out, however, that the two sampling procedures (random sampling and pedagogical sampling) will not always yield equivalently informative samples, and that in general, use of pedagogical sampling by a teacher (and an assumption of pedagogical sampling on the part of the learner) yields more efficient learning. For example, teachers engaging in pedagogical sampling may emphasize exemplars that clearly communicate the boundaries of a distribution, or are particularly informative for ruling out alternate hypotheses (e.g., a teacher might purposefully select to demonstrate that whales have lungs to show that all mammals have them, because this would rule out the alternate hypothesis that only land-mammals have lungs; in contrast, randomly sampling from mammals might yield a sample of only land-mammals).

The present studies, along with prior work, suggest that attention to sample composition in pedagogical and non-pedagogical learning is supported by different cognitive processes and follows different developmental trajectories. Although 5-year-olds clearly possess the information-processing skills to reason about sample composition (as demonstrated by their performance in pedagogical contexts), they appear not to attend to sample composition in non-pedagogical contexts until later in childhood. Thus, these findings highlight the need to consider not only the *evidence* that children are presented with, but also how the learning context influences children's interpretation of and use of evidence for knowledge development.

Acknowledgments

This research was supported by NICHD grant HD-36043 to the second author. We are grateful to the children, teachers, and parents of Abbot Elementary School, Ann Arbor Open Elementary School, and Cornerstone Elementary School for their participation in this research. We also thank Christine Shenouda and Holly Keilch for their assistance with data collection, and Tamar Kushnir for helpful input on this work.

References

Carey, S. Conceptual change in childhood. Cambridge, MA: Bradford Books; 1985.

Csibra, G.; Gergely, G. Social learning and social cognition: The case for pedagogy. In: Munakata, Y.; Johnson, MH., editors. Processes of Change in Brain and Cognitive Development. Attention and Performance XXI. Oxford: Oxford University Press; 2006. p. 249-274.

Ellis, AT.; Denison, SM.; Xu, F. Camshafts and chlorophyll: Statistical information and category-based induction in preschoolers; Poster presented at the biennial meeting of the Cognitive Development Society; Santa Fe, New Mexico. 2007.

Feeney, A.; Heit, E., editors. Inductive reasoning: Experimental, developmental, and computational approaches. New York: Cambridge University Press; 2007.

Gelman SA. The development of induction with natural kind and artifact categories. Cognitive Psychology. 1988; 20:65–95. [PubMed: 3338268]

Gelman, SA. The essential child: Origins of essentialism in everyday life. New York: Oxford University Press; 2003.

Gergely G, Egyed K, Király I. On pedagogy. Developmental Science. 2007; 10:139–146. [PubMed: 17181712]

Gutheil G, Gelman SA. Children's use of sample size and diversity information within basic-level categories. Journal of Experimental Child Psychology. 1997; 64:159–174. [PubMed: 9120379]

Heit E. Properties of inductive reasoning. Psychonomic Bulletin & Review. 2000; 7:569–592. [PubMed: 11206199]

Heit E, Hahn U. Diversity-based reasoning in children. Cognitive Psychology. 2001; 43:243–273. [PubMed: 11741343]

Heit, E.; Hahn, U.; Feeney, A. Defending diversity. In: Ahn, W.; Goldstone, RL.; Love, BC.; Markman, AB.; Wolff, P., editors. Categorization inside and outside the lab: Festschrift in honor of Douglas L. Medin. Washington, DC: American Psychological Association; 2004.

Kushnir T, Wellman HM, Gelman SA. The role of preschoolers' social understanding in evaluating the informativeness of causal interventions. Cognition. 2008; 107:1084–1092. [PubMed: 18039543]

Lopez A. The diversity principle in the testing of arguments. Memory & Cognition. 1995; 23:374–382.

Lopez A, Gelman SA, Gutheil G, Smith EE. The development of category-based induction. Child Development. 1992; 63:1070–1090. [PubMed: 1446543]

Medin DL, Coley JD, Storms G, Hayes BL. A relevance theory of induction. Psychonomic Bulletin & Review. 2003; 3:317–332.

Osherson DN, Smith EE, Wilkie O, Lopez A, Shafir E. Category-based induction. Psychological Review. 1990; 97:185–200.

Rhodes M, Brickman D, Gelman SA. Sample diversity and premise typicality in inductive reasoning: Evidence for developmental change. Cognition. 2008; 108:543–556. [PubMed: 18436200]

Rhodes M, Gelman SA, Brickman D. Developmental changes in the consideration of sample diversity in inductive reasoning. Journal of Cognition and Development. 2008; 9:112–143.

Shafto, P.; Goodman, N. Teaching games: Statistical sampling assumptions for learning in pedagogical situations. In: Sloutsky, V.; Love, B.; McRae, K., editors. Proceedings of the 30th Annual Cognitive Science Society; Austin, TX: Cognitive Science Society; 2008.

Shipley EF, Shepperson B. Test sample selection by preschool children: Honoring Diversity. Memory & Cognition. 2006; 34:1444–1451.

Xu, F. Rational statistical inference and cognitive development. In: Carruthers, P.; Laurence, S.; Stich, S., editors. The innate mind: Foundations and the future. Vol. Vol. 3. New York: Oxford University Press; 2007.

Xu F, Garcia V. Intuitive statistics by 8-month-old infants. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105:5012–5015. [PubMed: 18378901]

Xu F, Tenenbaum JB. Word learning as Bayesian inference. Psychological Review. 2007a; 114:245–272. [PubMed: 17500627]

Xu F, Tenenbaum JB. Sensitivity to sampling in Bayesian word learning. Developmental Science. 2007b; 10:288–297. [PubMed: 17444970]

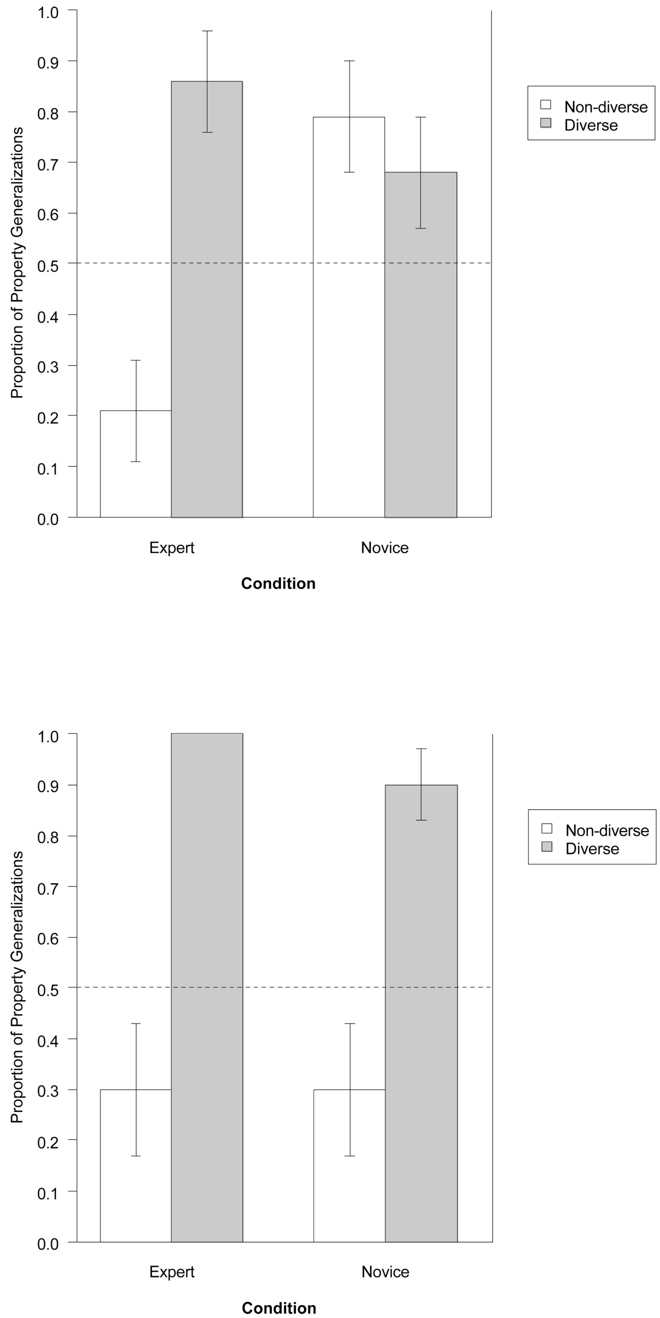

Figure 1.
Proportions of property generalizations to basic-level matches, by condition and sample, for five-year-olds (1a) and adults (1b), Study 1.

NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Description of visual materials, Study 1

Contents of a Sample Visual Display		
Dalmatian[*,^]	Dalmatian[*]	Dalmatian[*]
Golden Retriever[^]	Golden Retriever	Golden Retriever
Black Labrador[^]	Black Labrador	Black Labrador
Collie	Collie	Collie
Basset Hound	Basset Hound	Basset Hound
Monkey	Monkey	Monkey
Turtle	Turtle	Turtle
Test Questions		
1. Dalmatian	2. Monkey	3. Collie
4. Dalmatian	5. Basset Hound	6. Turtle

[*] *Note.* Items marked with a were included in the non-diverse teaching set.

[^] Items marked with a were included in the diverse teaching set. Exemplars were photographs of real animals. All exemplars from within the same subordinate class were noticeably different exemplars (e.g., they were different sizes, shades, had different markings). Test questions were presented individually.

Table 2

Individual Response Patterns, Study 1

	Subordinate-only	Basic-level	Other
CHILDREN			
Expert			
Non-diverse	9	2	2
Diverse	0	11	3
Novice			
Non-diverse	3	9	2
Diverse	2	9	3
ADULTS			
Expert			
Non-diverse	5	2	3
Diverse	0	10	0
Novice			
Non-diverse	6	2	2
Diverse	0	8	2