Chapter to appear in: *Advances in Child Development and Behavior: Rational Constructivism in Cognitive Development*

Irrational Inferences?

When children ignore evidence in category-based induction

Marjorie Rhodes

Department of Psychology, New York University

6 Washington Place, room 301, New York, New York, 10003; 212-998-3546;

marjorie.rhodes@nyu.edu.

Key Words: categorization, induction, folkbiology, rational inference, diversity, conceptual development

Table of Contents

Abstract

The process of induction—generalizing information obtained from limited samples to inform broader understandings —plays a critical role in learning across the lifespan. Previous research on the development of induction has found important developmental changes in one critical component of induction—how children and adults evaluate whether a sample of evidence is informative about a broader category. In particular, when acquiring knowledge about biological kinds, adults view samples that provide diverse representation of a category (e.g., an eagle, penguin, and robin, for the category *birds*) as more informative than a less diverse sample (e.g., three robins) for drawing inferences about the kind. In contrast, children under the age of eight often neglect this feature of sample composition, viewing both types of samples as equivalently informative. Is this a case of children making irrational inferences? This chapter examines how these findings can be reconciled with rational constructivist approaches to cognitive development, focusing on (1) the role of the sampling context in determining how learners incorporate information about sample composition into inductive inferences, and (2) how developmental differences in learners' intuitive theories influence how they make sense of new evidence. This chapter highlights how strong tests of rational approaches come from incidences where children's performance appears to be quite non-normative.

Irrational Inferences?

When children ignore evidence in category-based induction

I. Introduction

The theme of this volume is that children are rational constructivists; that is, that they actively try to make sense of their environment (the constructivist part) and that they do so by integrating new evidence with their prior beliefs via a process that approximates probabilistic statistical inference (the rational part; Xu, 2007). As shown by the numerous chapters in this volume, as well as by the breadth of topics they cover, this perspective has led to advances in our understanding of cognitive development across a wide range of questions. A rational constructivist account of cognitive development is compelling because it describes how domain-general learning mechanisms can give rise to domain-specific knowledge, applies across a wide range of learning tasks, and moves beyond old arguments between nativism and empiricism to describe a learning process that both accounts for prior knowledge and allows for change (Gopnik & Wellman, in press).

A central question in developmental psychology is whether the mechanisms by which we acquire knowledge are similar across development or undergo fundamental change. The answer to this question from the perspective of rational constructivism is clear: Although there may be substantial changes in children's knowledge and theories across development, the learning process—how beliefs are updated in response to new evidence—relies on similar mechanisms from early infancy (Denison, Reed, & Xu, in press; Dewar & Xu, 2010) to childhood (Xu & Tenenbaum, 2007a, 2007b; Schuz, Bonawitz, & Griffiths, 2007) through adulthood (Kemp & Tenenbaum, 2009; Tenenbaum, Kemp, & Shafto, 2007). This perspective challenges us to think critically about instances when children's responses to evidence are quite different from adult responses, and in particular, instances where children's inferences appear—at least at first, second, and even third glance—to be irrational.

In this chapter, I examine an instance of apparent irrationality, focusing on how children acquire generic knowledge about biological kinds. Generic knowledge is generalizable knowledge about abstract kinds. For example, generic knowledge about dogs includes that they bark, have four legs, and have fur. Generic knowledge is not tied to specific individuals and need not apply to every individual category member. Yet, because kinds cannot be directly observed, much generic knowledge is acquired via induction from individual instances (Rips, 1975). For example, upon learning that a particular dog barks, a child might infer that dogs—in general—bark. This facilitates use of generic knowledge to make predictions about new instances. For example, upon learning that a single bird has hollow bones, we can assume that other birds will have them; upon experiencing that a birthday party involves bringing presents and eating cake, we can predict what will take place at other parties in the future; upon learning that one girl likes sparkly stickers, we might expect other girls to like sparkly stickers as well. As demonstrated by these examples, these inferences are probabilistic and not always accurate. Yet, the underlying mechanisms reflected here—categorizing individuals as fundamentally alike and using these categories to generalize knowledge—are fundamental means by which we make sense of, predict, and interact with the environment. Although the acquisition of generic knowledge via induction is an important process across many domains, here I focus on these issues particularly for biological kinds. This focus follows much previous work on induction, and is of particular interest because of the rich, hierarchical structure of the biological world (Atran, 1990).

II. The Problem of Induction and a Rational Solution

Imagine that a child encounters a dog for the first time (perhaps, her friend's golden retriever), and learns that he is friendly and likes to be petted. How far should she generalize this knowledge? Certainly, it seems reasonable to generalize this knowledge to other encounters with this particular dog. Thus, if she encounters the dog again the next day, she can confidently go right up and pet him. But, how much farther should she generalize this knowledge? Should she now assume that other golden retrievers are friendly too, so that she can confidently approach and pet other golden retrievers that she meets? What about other dogs more generally? Or other four-legged furry animals? Clearly, some generalization beyond the individual is warranted—assuming that other golden retrievers will be friendly is a pretty good bet. But, generalizing too far could get the child into trouble; if she generalizes to all dogs, she could find herself in a dangerous situation if she encounters a vicious pit bull, and if she generalizes to all four-legged furry animals, even more so if she encounters a mountain lion.

Determining how far to generalize information is a basic problem of all inductive inference. Rational constructivist approaches suggest one perspective on how such inferences might be constrained. For example, Xu and Tenenbaum (2007a, 2007b) demonstrated that children are attentive to the composition of evidence within a sample, can infer the population that particular samples represent, and then restrict their generalizations to the represented population. They examined these processes in a word-learning context, another key inductive challenge of childhood. On their task, children were shown an array of different kinds of dogs, for example, including a range of sub-types (golden retrievers, basset hounds, dalmations, and so on), and were shown a sample of three of these dogs labeled as “blickets”. On this task, children attended closely to the sample composition to determine how far to extend the word blicket on subsequent test items. If children were shown three golden retrievers labeled as blickets, they extended the term only to other golden retrievers. In contrast, if they were shown three dogs of different sub-types, they extended the word to the basic level.

From a rational perspective, children’s performance in this task can be understood as follows. First, when shown the very first piece of evidence, children might consider multiple hypotheses: (a) that blicket is the name of the sub-type “golden retriever” or (b) that blicket is the name of the basic level category “dog”. Perhaps they consider these two hypotheses to be approximately equally likely, before they are exposed to further evidence. Second, after they are exposed to further evidence (e.g., two more golden retrievers, or two different dogs) they evaluate the likelihood that they would have seen each set of evidence, under each of the initial hypotheses. Whereas both sets of evidence are consistent with the hypothesis that blicket refers to “dog,” the observation of a sample of all golden retrievers under this hypothesis is unlikely. Xu and Tenenbaum suggest that children would view this sample as a “suspicious coincidence.”

In other words, the likelihood of the obtained sample is higher under the hypothesis that blicket refers to golden retriever than the hypothesis that blicket refers to dog, leading children to favor the narrower hypothesis in this case.

The extension from this word-learning example to the acquisition of generic knowledge ought to be straightforward. For example, if the child observed that three golden retrievers are friendly and like to be petted, she should feel confident that the property “is friendly” applies to golden retrievers, but less sure that it applies to dogs generally. In contrast, if she observed that three dogs of different subtypes are friendly and like to be petted, she should extend the property to the basic level. Extensive prior work suggests, however, that in just these types of contexts, children’s inferences often appear to neglect these dimensions of sample composition. Instead, children have been found to generalize to basic level categories equally often based on less representative evidence (the three golden-retrievers, hereafter, the “non-diverse sample”) and more representative samples (the three different dogs, hereafter, the “diverse sample”), as reviewed below. Does this neglect of sample composition in children’s reasoning reflect irrationality? Does this indicate, for example, that they fail to update the probabilities of their initial hypotheses following differing samples of evidence? To begin to address these questions, I first review the evidence that children often fail to consider sample composition when acquiring generic knowledge about biological kinds via induction.

III. Evidence that Children have Difficulty Being Rational

Extensive prior work shows that adults’ are very sensitive to sample composition for induction involving biological kinds. In particular, adults view samples that provide diverse representation of a category (as in the 3 dogs of different sub-types above) as supporting conclusions about the kind (dogs), but samples that provide less diverse representation of a category (as in the 3 golden retriever sample above) as supporting narrower inferences. For example, as reviewed below, adults view a diverse sample (e.g., a penguin, owl, and robin) as more informative than a non-diverse sample (e.g., three robins) for drawing inferences about abstract categories (e.g., birds). These effects have been termed “diversity effects,” and are a primary feature of adult inductive inference that models of induction have aimed to explain (Heit, 2000).

Adults robustly consider sample diversity when evaluating samples of evidence. Osherson et al. (1990) found that adults rated inductive arguments that built on diverse premises as stronger than inductive arguments that built on less diverse premises. For example, an argument supporting a general conclusion (e.g., “all mammals have four-chamber hearts”) was rated as stronger when it built on premises that included a diverse sample of mammals (e.g., “dogs and whales have four-chamber hearts”) rather than a less diverse sample of mammals (e.g., “dogs and wolves have four-chamber hearts”). Also, evidence selection tasks show that adults seek out diverse evidence. For example, to test if a property is true of mammals, adults prefer to examine diverse over non-diverse samples of mammals (Kim & Keil, 2003; Lopez, 1995; Lopez et al., 1997; Rhodes, Gelman, & Brickman, 2008; Rhodes, Brickman, & Gelman, 2008). Sample diversity is not the only criterion that adults use to evaluate evidence—they can also rely on causal knowledge, for example, or other specific background knowledge that is relevant to the question at hand (Medin, Coley, Storms, & Hayes, 2003). Yet, sample diversity is an important factor that adults consider across many different types of experimental tasks (Heit, 2000; Lopez, Atran, Coley, Medin, & Smith, 1997). Critically, basing broader inferences on diverse samples is consistent with computational accounts of rational statistical inference (Heit, 2000; Tenenbaum et al., 2007).

Yet, in contrast to adults' robust consideration of sample diversity, young children have been found to neglect sample diversity across many inductive reasoning tasks. In many previous studies, children below the age of 9 have treated diverse and non-diverse samples as equivalently informative for inferences about kinds. For example, children between the ages of 5 and 8 are just as likely to infer that "all birds have hollow bones" after learning that three non-diverse birds have hollow bones (e.g., three robins) as learning that three diverse birds have hollow bones (e.g., an eagle, owl, and robin; Gutheil & Gelman, 1997; Lopez, Gelman, Gutheil, & Smith, 1992; see also, Li, Bihau, Li, Li, & Deak, 2009). Younger children also do not consider sample diversity when selecting evidence; for example, 6-year-olds are just as likely to select a non-diverse sample (e.g., two Dalmatians) as a diverse sample (e.g., a Dalmatian and a pit bull) to find out if something is true of a category (e.g., dogs; Rhodes, Gelman, et al., 2008, Rhodes, Brickman, et al. 2008). In these studies, diversity effects have been found to emerge around ages 8-9 and to become more robust by ages 10-11.

Children's failure to recognize the informative value of diverse evidence does not relate to difficulty with the task demands in these experiments, or to general problems noticing or processing information about sample diversity. For example, Rhodes, Gelman, et al. (2008) showed that although 6-year-olds could reliably distinguish diverse from non-diverse sets, they still did not prefer to base generalizations on information obtained from diverse samples. Also, Rhodes, Brickman et al. (2008) showed that 6-year-olds do apply some systematic strategies to selecting between samples on the types of questions used in these experiments. For example, 6-year-olds preferred to base inferences on samples containing typical, over atypical, exemplars. Yet, although children showed some systematic criteria for evaluating samples, they once again did not factor sample diversity into these judgments.

More generally, Heit & Hahn (2001) and Shipley & Shepperson (2006) provide evidence that children can distinguish diverse from non-diverse samples and reason about them systematically. For example, Heit and Hahn presented 5-year-olds with two samples of dolls, including three diverse dolls belonging to one character and three non-diverse dolls belonging to another character. Participants were then shown a different doll, and were asked to predict to whom the doll belonged. On these questions, children reliably responded that the target toy belonged to the character that owned the diverse set of toys. These findings demonstrate that young children can recognize sample diversity and sort based on diversity. They do not, however, indicate that young children view diverse samples as a stronger basis for induction. Instead, children may solve these problems by recognizing that diverse items better match diverse (rather than non-diverse) sets of evidence. Similarly, Shipley & Shepperson (2006) report that preschool children prefer to test toys from two sub-classes (e.g., one blue whistle and one red whistle) in order to determine if whistles "make good party favors." Although this task also reveals that young children recognize and reason about sample diversity, because participants were not asked to make an inference about a larger set (e.g., whistles not included in either specific sub-class), this study also does not provide evidence that young children use sample diversity to determine whether a sample provides a good basis for broader generalizations (i.e., about a larger category or unobserved instance).

Thus, children's failure to consider sample composition when evaluating samples of evidence does not relate to task demands, difficulty noticing or processing diversity, or more general difficulty making decisions based on sample diversity. As shown by the work by Xu and Tenenbaum (2007a, b) discussed above, as well as other recent work, it is also clear that children's failure to consider sample diversity does not relate to general problems in statistical reasoning, or in particular, reasoning about how samples represent populations. Xu and colleagues have shown that quite sophisticated abilities reasoning about the relation between samples and populations emerge early in infancy. For example, infants expect a sample drawn from a bowl containing equal proportions of blue and yellow balls to contain equal proportions of blue and yellow, and are surprised if the obtained sample is all blue (Xu & Denison, 2009; Xu & Garcia, 2008). Infants are not applying a simple matching strategy to solve these tasks; their expectations break down if they know the sampler has a preference for blue, for example, suggesting that they only expect samples to accurately represent populations if they are drawn randomly.

Given these statistical abilities for reasoning about the relation of samples to populations, it is surprising that children do not consider sample composition when acquiring new generic knowledge via inductive reasoning. Children ought to realize, for example, that a diverse sample of dogs provides better representation of the category dogs than a non-diverse sample. Why do children fail to consider sample composition during category-based induction? Does this failure indicate that the mechanisms that support the acquisition of generic knowledge in childhood are somehow less rational than those that do so in adulthood? To tackle this issue and consider its implications for cognitive development, I next consider two factors that may enable children to consider sample composition in a rational manner, towards the aim of better understanding the circumstances when they do not.

IV. Two Factors that (may) Influence Rationality

A. The Learning Context

A critical difference between the category-based induction studies reviewed above and Xu and Tenenbaum (2007a, 2007b) is the *purpose* of the sample being presented. In Xu and Tenenbaum, an intentional adult selects the samples in order to show the child a new concept. Thus, the goal of the sample is information *communication*—the adult has some information (e.g., the meaning of the word blicket) that the child does not have, and the adult selects a sample to demonstrate the concept. In a broad sense, this interaction represents a pedagogical exchange. From this perspective, pedagogical learning does not require a formal teacher, but instead is defined by the epistemic gap between the teacher and the learner, as well as the intent of the teacher to communicate information to the learner. In contrast, in the induction experiments summarized above, children were not taught properties on purposefully presented samples; rather they were asked to *discover* new information via induction.

Thus, one possibility is that sample composition plays different conceptual roles in the *discovery* versus *communication* of new knowledge.

During information communication (teaching), 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 kinds of 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) show that adults assigned to teaching roles readily engage in this kind of effective sampling—without explicit instruction to do so—picking samples that will most unambiguously and efficiently communicate the underlying distribution to the learner. Further, adults assigned to learner roles assumed that teachers would provide informative samples, and this assumption helped them learn more efficiently.

During information communication, sample composition provides a window into the communicative goal of the teacher. For example, a learner may assume, “The teacher is trying to teach me something. She had all these dogs to choose from, but she chose three of the same kind. That decision was purposeful and intended to help me learn. She must be trying to tell me that this is about just that 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. In this manner, the “suspicious coincidence” noticed by the children in Xu and Tenenbaum, (2007a, 2007b) is suspicious with respect to the adult’s intent (e.g., “if she were trying to teach me about all dogs, it would be odd for her to pick this narrow sample”) not to the state of the world (e.g., *not* “if this were true of all dogs, it would be odd that I’ve only encountered this narrow sample).

In contrast, when learning involves knowledge *discovery* instead of *communication*, sample composition is relevant to testing hypothesis about the world. A diverse sample of dogs provides a strong test of a hypothesis about dogs as a kind, for example, whereas a non-diverse sample of dogs provides a weak test (Heit, Hahn, & Feeney, 2005). Thus, one possibility is that children recognize sample composition as a window into the communicative intent of a teacher—and thus consider sample composition in a rational manner in these cases—but not as an indicator of how informative a sample is for the process of information discovery.

Rhodes, Gelman, and Brickman (2010) directly tested this hypothesis, using a method similar to Xu & Tenenbaum (2007a). In this work, 5-year-old children and adults were exposed to an array of animals from a basic level category (e.g., an array of dogs). Participants were shown samples to help them learn a new fact about the animal category (e.g., to find out which animals have an epithelium inside). Across condition, samples were presented either by an animal expert who knew a great deal about the animals, or a novice who did not know anything about them (see Kushnir, Wellman, & Gelman, 2008). In the Expert condition, across items, the expert either presented a diverse sample of dogs (a basset hound, dalmation, and golden retriever) or a non-diverse sample of dogs (three basset hounds), with the clearly stated intention of teaching “which animals have an epithelium inside”. In the Novice condition, across items, the puppet checked various animals (the same exemplars as were shown by the Expert), to discover if they had an epithelium and reported the results to the child. Because the novice did not know which animals had the property ahead of time, it should have been clear to children that they were not selected with particular communicative intent. The aim of the Expert condition (from the child’s perspective) was information communication, whereas the aim of the Novice condition was to discover the information along with the puppet. Across conditions, children were exposed to identical evidence presented by a puppet; the conditions varied only in whether the sample was systematically selected by a knowledgeable teacher to communicate information or by a novice aiming to discover information.

Indeed, in the Expert condition, preschool-age children inferred that the property applied only to the subordinate category (e.g., only to dalmations) when they were shown a non-diverse sample, but to the basic level kind when they were shown a diverse sample. In contrast, in the Novice condition, 5-year-olds extended the property to the basic level following both types of samples. Adults, in contrast, showed the same pattern across both conditions—they reliably extended to the subordinate following the non-diverse sample and to the basic level following diverse samples, regardless of who presented the sample. These data suggest that preschoolers consider sample composition during learning events where the learning goal involves information *communication,* but not information *discovery*. In contrast, adults consider sample composition in both types of learning contexts.

In a follow-up study, Rhodes et al. further compared 5-year-olds understanding of sample composition for knowledge *communication* vs. *discovery* by placing the children themselves either in the position of "teacher" or "scientist". Here, children were asked to select samples either to *teach* someone else that a basic level category contains a novel property (e.g., that all dogs have an epithelium inside) or to *discover whether* a basic level category contains a novel property (e.g., whether all dogs have an epithelium inside). Children were offered a choice between diverse and non-diverse samples of dogs. In the teacher condition, children indeed reliably selected the sample that provided diverse representation of the category, whereas in the scientist condition, children responded at chance. These data provide further evidence that children recognize the role of sample composition in effectively communicating information, but not as an indicator of the strength of samples for hypothesis testing.

These studies help to resolve the apparent discrepancy between Xu and Tenenbaum and studies of category-based induction, by suggesting that children recognize the role of sample composition in information communication prior to information discovery. Yet, they do not address why children fail to consider sample composition during information discovery on these tasks. Although showing that children engage in rational inference in pedagogical contexts in an important step to characterizing how children acquire generic knowledge, much knowledge acquisition occurs in the absence of knowledgeable teachers who purposefully select samples for children. Thus, it is critical to examine why children fail to consider sample composition when they discover samples of evidence on their own, with an aim of resolving this pattern with the general rational constructivist framework.

B. Intuitive Theories

To consider why children fail to overlook sample composition when selecting their own samples to discover new generic knowledge, it is useful to consider their intuitive theories of the biological world, which ought to shape the types of evidence they view as relevant to these problems. Towards this aim, Rhodes and Brickman (2010) proposed that children have abstract expectations that biological kinds are highly homogenous (see Atran, 1990) and that these expectations lead them to treat diverse and non-diverse samples from a category as inter-changeable.

Expectations of category homogeneity entail the extent to which people assume that—despite superficial differences—all members of a category are fundamentally alike. There is abundant indirect evidence that young children expect some categories to be more homogeneous than adults do. For example, young children (ages 4-7) have strong expectations that the members of natural kind categories will demonstrate category-typical properties—even in the face of contrasting individuating information—whereas older children (age 10) and adults allow for more individual variation (Berndt & Heller, 1986; Taylor, 1996; Taylor, Rhodes, & Gelman, 2009). Also, preschool-age children are more likely than adults to believe that categories are objective and coherent (Kalish, 1998; Rhodes & Gelman, 2009), and to infer that a property observed in one individual will be found in other members of a kind (Gelman, 1988; Rhodes & Gelman, 2008). Young children also often neglect subtypes within basic level categories (e.g., they fail to recognize basset hounds and Dalmatians as meaningfully different kinds of dogs, Waxman, Lynch, Casey, & Baer, 1997), perhaps because they view basic level categories as highly coherent.

Gelman (2003) has argued that cognitive biases to assume that basic-level categories are homogeneous play a powerful role in early conceptual development and propel knowledge acquisition by allowing children to overlook superficial differences and focus on underlying regularities (e.g., in order to learn the conceptual category *dog*, children must overlook superficial difference in size and color and focus on the properties that all dogs share). Yet, focusing on similarities may also lead children to overlook meaningful and important variation (Gelman & Kalish, 1993). From this perspective, an important component of conceptual change across childhood entails increased consideration of within-category variability.

How might developmental differences in expectations about within-category homogeneity and variability influence how children and adults evaluate samples during inductive reasoning? Rhodes and Brickman (2010) proposed that children's strong expectations that biological categories are homogeneous lead them to be less discriminating about whether a given sample of evidence is informative (e.g., because children assume that all dogs are fundamentally alike, it does not matter to them which particular dogs are observed to support an inference about the category as a whole; Rhodes et al., 2008a, 2010).

To test this hypothesis, Rhodes and Brickman (2010) showed 7-year-olds and adults a set of perceptually and taxonomically diverse birds. Children assigned to a Variability condition were prompted to consider differences among birds (e.g., that some fly and some do not fly, some hunt for food and some dig for food, and so on). Children assigned to a Similarity condition saw the same visual stimuli, but were prompted to consider similarities (e.g., that all birds have feathers, all birds feed their babies mashed up foods, and so on). Children assigned to a Control condition were shown the same visual stimuli, but were not prompted to think about any properties. Following the primes, children completed measures of their consideration of sample diversity in evaluating samples of evidence. For example, they were asked to select between examining diverse samples (e.g., a robin and a blue jay) or non-diverse samples (e.g., two robins) to test whether a novel property is true of a category (e.g., "to find out if birds have gizzards inside"). Children in the Variability condition reliably chose diverse samples, whereas children in the Similarity and Control conditions performed at chance. Adults selected diverse samples in all conditions. This study experimentally demonstrates that increasing attention to within-category variability increases diversity-based reasoning among children.

The effect of the Variability primes was quite robust. The Variability primes improved performance on multiple measures of diversity-based reasoning and increased diversity-based reasoning for both the animal categories presented in the prime (e.g., birds) and for other animal categories (e.g., pigs, frogs). Thus, the prime appeared to function not by increasing children's specific knowledge about *birds,* but by challenging children's more generalized expectations about the homogeneity of animal categories. The effect of the Variability primes was also appropriately selective, however. A follow-up control study documented that although Variability primes increased diverse sample selections for inferences about broad categories (e.g., birds), they did not do so for inferences about specific sub-types (e.g., robins), for which picking a non-diverse sample (e.g., two robins) would be more informative. Thus, the Variability primes did not lead children to view diversity as better across the board, but rather, to engage in diversity-based reasoning in a normative manner.

The data reviewed above suggest that young children overlook sample composition when acquiring new generic knowledge about biological kinds because they have strong expectations that biological categories are homogenous. Is this neglect of sample composition irrational? Recent models (Kemp & Tenenbaum, 2009; Tenenbaum et al., 2007) suggest a possible way to reconcile this finding with a rational constructivist perspective. These models indicate that domain-specific intuitive theories shape the prior probabilities that people bring to learning events. As in all rational inference, from this perspective, inferences result from the interaction between participants' prior expectations (e.g., their prior estimate of the probability that a property found in one bird will be found in all birds, for example) and the new evidence that they receive. As described by Kemp and Tenenbaum (2009), these prior expectations can stem from intuitive theories or conceptual biases. Thus, developmental differences in these prior probability estimates (with younger children having higher baseline prevalence estimates, reflecting their assumptions that all category members are fundamentally alike) could explain developmental differences in sensitivity to sample diversity.

Yet, this possibility requires direct examination. This framework suggests that developmental differences in consideration of sample composition relate to differences in initial theories (i.e., prior expectations about how properties are distributed across categories), not differences in learning mechanisms. If so, although children initially view a hypothesis that a property applies to a basic level as more likely than a hypothesis that a property applies only to a subtype, they should be able to update their beliefs in response to new evidence if that new evidence is compelling enough. Yet, direct evidence that they do so is needed.

V. Rational in the end?

Mechanisms for rational statistical inference are clearly in place early in childhood. The data summarized above, as well as in other chapters of this volume, indicate that these mechanisms contribute to learning across a wide-range of content domains and learning challenges. In the present context, they appear to guide how children acquire generic knowledge via communication from experts.

So, are children's responses to evidence as they acquire generic knowledge via induction rational in the end? This remains an open question. During information communication, children use sample composition to constrain their inferences, consistent with rational models. Yet, whether they do so for information discovery—when they select samples themselves, when samples are produced via procedures that lack transparency, or when samples are produced in the absence of communicative goals—remains much less clear. These issues will be important to explore in future work, both to consider how learning mechanisms vary across pedagogical and non-pedagogical settings and to determine the scope of conceptual learning that can be accounted for by a rational constructivist perspective.

Whereas much of the work on rational constructivism has focused on identifying developmental continuities, rational constructivism also provides a useful framework for considering developmental differences. This perspective prompts us to consider such differences very carefully, and to determine whether differences reflect changes in the intuitive theories or in the mechanisms by which beliefs are updated in response to new evidence. The example of category-based induction discussed in this chapter illustrates that a key test of these models may lie in the cases where children's responses to evidence appear quite different from adults'.

The research reviewed in this chapter also has implications for the role of pedagogical cues in the acquisition of generic knowledge. Some have argued that pedagogical cues serve to signal to children that information is generic (Csibra & Gergely, 2009). From this perspective, when children realize they are being taught by a knowledgeable teacher, they assume that the demonstrated information is generalizable. Although the present studies are consistent with the broad proposal that children respond differently to information in the presence of pedagogical cues, these data suggest an alternate conclusion regarding the particular effects of these cues. The data reviewed above suggest that children assume that information is generic in the *absence* of pedagogical cues. In particular, as shown in the Novice condition of Rhodes et al. (2010), in the absence of pedagogical cues, children generalized the information to basic level categories following both diverse and non-diverse samples. The Expert condition suggests that pedagogical cues functioned to help children properly *restrict* their inferences, not to broaden them. Similarly, Rhodes and Brickman (2010) indicated that children have strong assumptions of category homogeneity (that the members of categories share many generic features). These data are consistent with the proposal that children treat certain information as generic by default (Cimpian & Erickson, 2012). From this perspective, pedagogical cues are not necessary for children to treat information as generic, but rather to guide them to when they should apply information more narrowly. More generally, pedagogical cues may not signal either generic or specific information per se, but rather that a sample is being selected purposefully, and thus that children should pay attention to sample composition and generalize appropriately. Resolving the discrepancies across these perspectives is an important area for future work.

References

Atran, S. (1990). *Cognitive foundations of natural history: Towards an anthropology of science.* New York: Cambridge University Press.

Berndt, T. J., & Heller, K. A. (1986). Gender stereotypes and social inferences: A developmental study. *Journal of Personality and Social Psychology, 50*, 889–898.

Cimpian, A., & Erickson, L.C. (2012). Remembering kinds: New evidence that categories are privileged in children's thinking. *Cognitive Psychology, 64,* 161-185.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13,* 148-153.

Denison, S., Reed, C., & Xu, F. (in press). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-old infants. *Developmental Psychology.*

Dewar, S., & Xu, F. (2010) Induction, over hypothesis, and the origins of abstract knowledge: Evidence from 9-month-old infants. *Psychological Science, 21,* 1871-1877.

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20*, 65–96.

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday life*. New York: Oxford University Press.

Gelman, S. A., & Kalish, C. W. (1993). Categories and causality. In R. Pasnak & M. L. Howe (Eds.), *Emerging themes in cognitive development* (pp. 3–32). New York: Springer-Verlag.

Gopnik, A., & Wellman, H.M. (in press). Reconstructing constructivism: Causal models, Bayesian learning mechanisms and the theory theory. *Psychological Bulletin.*

Heit, E. (2000). Properties of inductive reasoning. *Psychonomic Bulletin & Review, 7*, 569–592.

Heit, E., & Hahn, U. (2001). Diversity-based reasoning in children. *Cognitive Psychology, 43,* 243–273.

Heit, E., Hahn, U., & Feeney, A. (2005). Defending diversity. In W. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. Wolff (Eds.), *Categorization inside and outside the lab: Festschrift in honor of Douglas L. Medin* (pp. 87–99). Washington, DC: American Psychological Association.

Kalish, C.W. (1998). Natural and artificial kinds: Are children realists or relativists about categories? *Developmental Psychology, 34,* 376-391.

Kemp, C., & Tenenbaum, J. (2009). Structured statistical models of inductive reasoning, *Psychological Review, 116,* 20-58.

Kim, N. S., & Keil, F. C. (2003). From symptoms to causes: Diversity effects in diagnostic reasoning. *Memory and Cognition, 31*, 155–165.

Kushnir, T., Wellman, H. M. & Gelman, S. A. (2008). The role of preschoolers' social understanding in evaluating the informativeness of causal interventions. *Cognition, 107,* 1084-1092.

Li, F., Bihua, C., Li, Y., Li, H., & Deak, G. (2009). The law of large numbers in children's diversity-based reasoning. *Thinking & Reasoning, 15,* 388-404.

Lopez, A. (1995). The diversity principle in the testing of arguments. *Memory and Cognition, 23*, 374–382.

Lopez, A., Atran, S., Coley, J. D., Medin, D. L., & Smith, E. E. (1997). The tree of life: Universal and cultural features of folkbiological taxonomies and inductions. *Cognitive Psychology, 32*, 251–295.

Lopez, A., Gelman, S. A., Gutheil, G., & Smith, E. E. (1992). The development of category-based induction. *Child Development, 63*, 1070–1090.

Medin, D. L., Coley, J. D., Storms, G., & Hayes, B. L. (2003). A relevance theory of induction. *Psychonomic Bulletin & Review, 3*, 317–332.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97,* 185–200.

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

Rhodes, M., & Brickman, D. (2010). The role of within-category variability in category-based induction: A developmental study. *Cognitive Science, 34,* 1561-1573.

Rhodes, M., Brickman, D., & Gelman, S.A. (2008). Sample diversity and premise typicality in inductive reasoning: Evidence for developmental change. *Cognition, 108,* 543-556.

Rhodes, M., & Gelman, S.A. (2008). Categories influence predictions about individual consistency. *Child Development, 79,* 1271-1288.

Rhodes, M. & Gelman, S.A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology, 59,* 294-274.

Rhodes, M., Gelman, S.A., & Brickman, D. (2008). Developmental changes in the consideration of sample diversity in inductive reasoning. *Journal of Cognition and Development, 9,* 112-143.

Rhodes, M., Gelman, S.A., & Brickman, D. (2010). Children's attention to sample composition in learning, teaching, and discovery. *Developmental Science, 13,* 421-429.

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

Shipley, E. F., & Shepperson, B. (2006). Test sample selection by preschool children: Honoring diversity. *Memory and Cognition, 34*, 1444–1451.

Taylor, M. (1996). The development of children's beliefs about social and biological aspects of gender differences. *Child Development, 67*, 1555–1571.

Taylor, M.G., Rhodes, M., & Gelman, S.A. (2009). Boys will be boys, cows will be cows: Children's essentialist reasoning about human gender and animal development. *Child Development, 79,* 1270-1287.

Tenenbaum, J., Kemp, C., & Shafto, P. (2007). Theory-based Bayesian models of inductive reasoning. In A. Feeney & E. Heit (Eds.), *Inductive reasoning:*

Experimental, developmental, and computational approaches. New York: Cambridge University Press.

Waxman, S., Lynch, E., Casey, L., & Baer, L.(1997). Setters and Samoyeds: The emergence of subordinate level categories as a basis for inductive inference. *Developmental Psychology, 33,* 1074-1090.

Xu, F. (2007). Rational statistical inference and cognitive development. In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The Innate mind: Foundations and Future, Vol. 3.* Oxford University Press.

Xu, F. & Denison, S. (2009) Statistical inference and sensitivity to sampling in 11-month-old infants. *Cognition, 112,* 97-104.

Xu, F. & Garcia, V. (2008) Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 5012-5015.

Xu, F., & Tenenbaum, J.B. (2007). Word learning as Bayesian inference. *Psychological Review, 114,* 245-272.

Xu, F., & Tenenbaum, J.B. (2007). Sensitivity to sampling in Bayesian word learning. *Developmental Science, 10,* 288-297.