CHAPTER FIFTEEN

Commentary on Part II
Abilities and Assumptions Underlying Conceptual Development

Ellen M. Markman and Vikram K. Jaswal

Some of the most fundamental questions one can ask about cognitive development are related to the nativist-empiricist debate: How do we best characterize the initial state of the infant, and how do we explain the transition from the initial state to the mature cognition of adults? Carey and Markman (1999) argued that, at heart, this is a debate about whether there are domain-specific learning mechanisms. The empiricist position is that the initial state can be characterized in terms of perception and that general, all-purpose learning mechanisms, which can form associations and detect correlations, are sufficient to account for the acquisition of knowledge. Many of the outstanding issues in cognitive development center on questions of the domain specificity of learning and of the need to postulate core concepts or constraints on learning. Also central to the field are questions of whether cognition, especially the cognition of infants and young children, should be characterized as perceptual or conceptual, and whether attributing "theories" to infants and young children is necessary to account for their understanding of the world or whether it is an anthropomorphic overinterpretation of the evidence. These six chapters on the nature of children's categories raise many of these most basic questions about human cognition and its development and provide thoughtful discussions, pertinent evidence, and elegant models for how best to characterize children's classifications. We will emphasize two of these questions in our comments: (1) Are children's classifications theory based or derived solely from computing correlations over perceptual features? (2) What role does language, in terms of naming categories, play?

Although we acknowledge that many of the questions remain open, we believe that the evidence favors a theory-based view of category development and suggests that language plays a critical role in the acquisition of a conceptual system, one that is not accounted for in an associationist model. There are two important qualifications to this conclusion, however. First, regardless of which position one takes, there is clear agreement that computing correlations over perceptual features plays an important role in children's developing categorizations, and that perceptual features are routinely used to identify category members. The debate is about whether such features and correlations are the whole story, not about whether they exist or are useful. Second, "theory-based" is not always an apt description. "Knowledge-based" may sometimes be closer, as might Mandler's (chap. 5) ideas about "interpretation," but even these notions often fail to capture the basis of children's classifications. We will argue that children often may rely on a vague, sometimes poorly specified and poorly understood notion that there is more than meets the eye, that there must be something that explains why things have the properties they do, but children may not know what it is.

This vague belief differs from children's naïve theories in domains such as physics, psychology, number, or space. In each of these domains, a theory is attributed to children because each of these conceptual frameworks has ontological concepts at its core, and some causal principles that operate in the domain to explain phenomena and generate predictions. For other classifications, however, it is not always clear what these core concepts and causal principles would be. In many cases, the ontology and causality may be derived from the foundational domain in which the classification is embedded; in fact, we argue that this is an important function of these domains. Once a lay biology has been set up, for example, then for any animal, one might expect that it has internal parts that perform some vital functions; that it will need nourishment to grow; and so on. So one fully legitimate way to think of categories as theory driven is that specific categories import the causal principles from the larger domain of which they are a part. In addition, as Gelman and Koenig and Gopnik and Nazzi mention in their chapters, children sometimes seem to hold essentialist views with minimal understanding of what the "essence" might be. Gopnik and Nazzi suggest that essentialism functions as a placeholder for some as-yet-unspecified causal mechanism. It is as if the child reasons, for example, "I know there is something that makes this a bird. Something must have caused it to have a beak, claws, and feathers, but I really don't know what." We propose that such "theory-based" categorization derives from children's faith that everything has a cause. Admittedly, the evidence for children's belief in determinism is not overwhelming, but see Bullock, Gelman, and Baillargeon (1982) for an argument that preschoolers hold such a belief.

Suppose children's essentialist "theory-based" beliefs are often little more than articles of faith—a belief that there must be a cause for the properties an object manifests, but not much understanding of what that cause might be. Would this shallow understanding trivialize the notion of theory-based categorization? We think not. In fact, this belief may serve as scaffolding for seeking out and representing relevant information. It could serve as a mechanism for motivating children to look beyond the obvious and as a way of reconciling disparities between superficial appearances and classifications when they arise.
This shallow understanding also helps explain an inconsistency in the literature—why children sometimes reveal a shape bias in studies of classification (e.g., Baldwin, 1992; Imai, Gentner, & Uchida, 1994) yet show clear abilities to treat nonsimilar objects as members of the same kind and to base their inductions on kind membership rather than perceptual similarity (e.g., Gelman & Markman, 1986). The inconsistency hinges on whether children are asked to determine the appropriate categorization on their own versus whether they are asked to learn and use a classification provided for them. On their own, with only a vague notion of the causal mechanism, they may not spontaneously construct a nonobvious classification. But—and this is a critical point—children are not limited to categories they can construct on their own: When category membership is stipulated for them through a label, they willingly form nonobvious categories. Thus, children make two basic assumptions that work synergistically to enable them to move well beyond the classifications they could construct on their own: (1) they believe in a deeper, nonperceptual basis for categorization, and (2) they are willing to accept labels for things that to them may not seem to be the same kind of thing.

With this framework in mind, we turn now to a brief commentary on each of the preceding six chapters.

Associationist Approaches

The chapters by Mareschal (chap. 14) and by Smith, Colunga, and Yoshida (chap. 11) provide explicit, detailed models to tackle important and controversial issues in category development. Both propose that low-level, atheoretical, associative mechanisms, capable of extracting statistical regularities from the input and environment, are sufficient to explain the development of children’s early categories (see also Younger, chap. 4). Links between features that frequently co-occur in the input are strengthened, while those between relatively infrequent, variable features are not. With time and training, certain expectations develop, so that on presentation of one feature (or more), a network (or child) will expect that other features that have been correlated with that one in the past will also be present.

Mareschal’s (chap. 14) approach is impressive for its ability to model interesting aspects of behavior, to generate novel predictions, and to then verify these predictions through clever experimental research. For example, using an habituation paradigm, Quinn, Eimas, and Rosenkrantz (1993) found that 3.5-month-old infants exposed to exemplars of cats formed a category of Cat that excluded novel dogs, but those exposed to exemplars of dogs formed a category of Dog that did not exclude novel cats. Mareschal and his colleagues developed a connectionist autoencoder that succeeded in simulating the same asymmetry by exploiting the fact that the perceptual variability in the Cat category was smaller than that in the Dog category (Mareschal, French, & Quinn, 2000). For children and the network alike, this difference in perceptual variability made novel dogs unlikely cats but novel cats plausible dogs. In fact, on the basis of the distribution of perceptual features in the categories, even when the network was familiarized with a mixed set exemplars containing more dogs than cats, Mareschal and colleagues predicted that the same asymmetry would emerge. Indeed, it did, both in their model and when they conducted a new study with 3- and 4-month-olds.

Mareschal also developed a network to model what he calls the “percept-to-label” shift, or a shift from reliance on perceptual information to reliance on a category label when making inductive inferences. The model starts out more attuned to perceptual information than label information because it is assumed that younger children have less experience with labels. Thus, when the young network is shown an image that yields one representation but is then informed that it has a label that provides a different representation, it weighs the representation provided by the percept more heavily than the one provided by the label. As the network is trained that certain labels go with certain percepts, however, it learns that the label input is a more reliable guide to an accurate representation than the percept alone, because there is more variability in the perceptual input.

The performance of the network on this task does simulate a perceptual-to-conceptual shift, a developmental phenomenon that Piaget (Inhelder & Piaget, 1964) and others have suggested. However, along with a number of authors in this book (Gelman & Koenig; Mandler; Oakes & Madole; Rakison), we are not convinced that there is such a shift in development. On our view, whether perceptual or conceptual information is used to make an inference depends on the inference to be made and the theories possessed by the child. For example, Mareschal suggests that 4-year-olds rely on category labels more than perceptual information when making inductions about natural kinds (Gelman & Markman, 1986) and argues that results with younger children have been more equivocal. In particular, he writes that 2-year-olds in a study by McCarrell and Callanan (1995) found “perceptual information to be a more robust basis for induction than object labels” (p. 372). It is true that 2-year-olds found the perceptual information more compelling than the object labels in McCarrell and Callanan’s study, but so did 4-year-olds. Indeed, even adults would probably find the perceptual information more compelling than the object labels using the materials from that study. For example, an animal with an accordion-like neck and called a wug was said to “stretch tall to eat cherries.” When shown another animal that shared the accordion-like neck but that was called by a different name, preschoolers were quite likely to agree that it “stretches tall to eat cherries,” and so would adults, we imagine. These results differ from those of Gelman and Markman because the stimuli were designed in radically different ways: McCarrell and Callanan’s stimuli were designed explicitly to provide obvious correspondences between an object’s form and function; Gelman and Markman’s stimuli were designed explicitly to avoid any such cues.

Mareschal (chap. 14) seems to argue that a theory-based account of category development rests on a total reliance on category labels in inductive inference tasks. In fact, however, an ability to use perceptual and label information flexibly seems to be the norm, and seems very much in line with theory-based approaches: people use a label on induction tasks when it is relevant to the inference, and not otherwise, and they decide what is relevant by drawing on theory-based knowledge. For example, it would be inappropriate to use a category label to make an inference about whether a baseball bat could fly like a vampire bat, or whether a computer mouse would like cheese like a field mouse (see also P. Bloom, 2000). These
extend a new word given to a novel complex solid object on the basis of its shape, and a new word given to a novel simple nonsolid on the basis of its material. However, whereas English-speaking children are also quite likely to extend a new word for a simple solid object on the basis of shape, Japanese-speaking children are more equivocal—some extend it on the basis of shape and others on the basis of material. Imai and Gentner argue that their results provide evidence for both prelinguistic, universal ontological knowledge about individuation (the similar results for complex objects and for nonsolids) as well as for more language-specific influences: English speakers group both simple and complex objects together as encoded by count nouns and as different from nonsolid objects, which are encoded as mass nouns; Japanese makes no grammatical distinctions between count and mass nouns, and so Japanese-speaking children can interpret a label applied to a simple object as a substance/material term or as an object term. In ambiguous cases then, the language being learned can influence children’s interpretations of new words (Gentner & Boroditsky, 2000).

Smith et al. (chap. 11) suggest that, in fact, both the “universals and the differences are the product of the same statistical learning mechanism” (p. 280). On their model, even the ontological knowledge about individuation common in English and Japanese is the result of correlations in the input. As their rating studies indicate, toddlers exposed to either language learn the same kinds of words, and those words encode the same kinds of correlated perceptual features (e.g., solid objects encoded by words are consistent in shape but not material; nonsolids encoded by words are consistent in material but not shape). This seems plausible for determining the mapping rules between ontologies and lexical categories, but there seems to be some confusion about whether we are considering ontologies only as evidenced in words or ontologies more generally. As Imai and Gentner (1997) have argued, their results do not bear on children’s understanding of the “essential properties and behavior of objects” (p. 193). Indeed, they argue that Japanese children in their study may have recognized that even simple objects had objectlike properties that would be different from the nonsolid stimuli (e.g., solidity, cohesiveness, etc.). What Imai and Gentner’s results (and Smith et al.’s model) bear on is the interesting and important question of how children learn to extend new words. It is an open question about how or even whether these results relate to how children think or reason about the thing to which the word applies.

Moreover, there is a fundamental difference between establishing an ontology in the first place versus simply mapping existing ontological categories onto grammatical form classes. There is good evidence that even very young infants who have no words in their productive vocabularies can distinguish between solids and nonsolids. For example, in one study by Huntley-Fenner, Carey, and Solimando (2001), 8-month-old infants watched as one solid object was placed behind a screen and then a second solid object was added. If the screen was removed to reveal a single object, infants looked longer than if the screen was removed to reveal two objects (see Wynn, 1992). On the other hand, when infants of the same age watched as one cupful of sand was placed behind the screen and then a second cupful was added, they looked equally long at displays consisting of a single small pile of sand or two small piles. In other words, they could track solid objects but not nonsolid ones. The basic onto-
The ontological distinction that Gelman and Koenig emphasize, and with good reason (e.g., see discussion by Mandler, chap. 5), is the animate-inanimate distinction. Perceptual features such as shape, texture, and eyes contribute to infants’ ability to distinguish animate from inanimate objects, but Gelman and Koenig cite evidence that babies and young children classify objects on the basis of animacy even when it is not based on such perceptual cues. The centrality of animacy is evident in the kinds of distinctions that languages make, and young children master a number of linguistic conventions that explicitly or implicitly code animacy. Gelman and Koenig remain uncertain about the age at which this ontological distinction emerges but argue that it has a clear privileged status by age 3.

Another implication of theory-based categorization is that causality should play a critical role in the representation of categories. Implicit causal knowledge guides attention to relevant features and their correlations in classification. Even in the face of salient noncausal perceptual features, very young children rely on features involved in causal relations to categorize objects and draw inductive inferences. We will return to the importance of causality when we discuss Gopnik and Nazzi’s chapter.

A theory-based representation of artifacts implies attention to functionally relevant over functionally irrelevant perceptual features. Gelman and Koenig summarize the controversy surrounding whether young children are capable of relying on function in classifying objects. In particular, Smith et al. (1996) claim that 3-year-olds are not influenced by function in naming objects while Kemler Nelson (Kemler Nelson, Russell, Duke, & Jones, 2000) and others find that young children do use function as a basis for extending object labels. Gelman and Koenig argue that differences in task demands explain the differences in findings, and that when function is made salient and comprehensible to young children, even 2-year-olds attend to and rely on functional cues in naming objects. Because form and function are often related, an understanding of function leads young children to attend to relevant perceptual features.

For animate objects, intention is a nonperceptual property that affects how very young children classify objects, the inferences they draw, and their behavior toward the objects. Woodward (1998) has provided striking evidence that even 6-month-olds construe actions performed by animate but not inanimate objects as goal directed. Thus, some of the conceptual ramifications of the ontological distinction between animate and inanimate things may be available even in young infants. Perceptual features, such as eyes, are doubtless important for this distinction and for attributing intentionality to objects. But, as Gelman and Koenig note, neither animacy nor intention can be reduced to perceptual features. For example, Johnson, Slaughter, and Carey (1998) found that 12-month-olds will follow the “gaze” of a creature without eyes if the object had earlier responded contingently to the baby.

Intention also plays a role in children’s reasoning about artifacts: children are sensitive to the function that the creator of an artifact had in mind when creating it and use that intended function in naming and reasoning tasks. For example, Bloom and Markson (1998) found that whether children considered a picture to be a lollipop or a balloon depended on their intent in drawing it, not on perceptual features that were the same in both cases. Similarly, Gelman and Bloom (2000) found that objects were considered to be artifacts if created intentionally rather than acciden-
tally. For example, children will judge a jagged piece of Plexiglas to be a knife if it was created through intentional rather than accidental means.

Many of the criteria already mentioned, such as the importance of function in classifying artifacts, of intention in distinguishing animate from inanimate objects, and of the intention of the creator in classifying drawings and artifacts, are evidence that children rely on nonobvious properties in constructing their classifications of objects. Gelman and Koenig cite a range of other studies that document children's weighing internal parts or substance in categorizing objects. Moreover, such reliance on internal parts occurs before children have detailed knowledge or understanding of the parts or their functions.

Another way in which children's categorizations are more than a record of correlated perceptual features is that children use categories to make predictions and draw inferences about novel properties of objects. By age 2½, children assume that categories provide inductive potential. As Gelman and Koenig argue, "children do not assume that labels are mere conveniences—ways of efficiently referring to perceptually encountered information in a shorthand way. Instead, children expect certain labels—and the categories to which they refer—to capture properties well beyond those they have already encountered" (pp. 344–345). Labels are a form of cultural transmission of knowledge—allowing one individual to efficiently pass on nonobvious information to other individuals and later generations—and may be critical in helping children form theory-based categories.

One major strength of the theory-based approach as outlined by Gelman and Koenig is the number of different sources of converging evidence that support their framework. One should ask, however, whether all of this evidence is equally strong and resistant to low-level explanations. That is, to what extent could an associationist mechanism account for the various pieces of evidence that Gelman and her colleagues have assembled? For example, could a pattern of correlations that includes both the perceptual and behavioral cues of an object be sufficient for detecting animacy? If lower-level, domain-general mechanisms could explain some of the evidence they cite, does this diminish the strength of their overall argument for theory-based categorization? Another issue we alluded to earlier concerns the extent to which children's categories really are theory based per se, instead of reflecting isolated facts, sketchy knowledge, or even simply a faith that there must be an underlying cause even though it is unspecified or poorly understood. In other words, there are different kinds of support or instantiations of "nonobvious" properties, ranging from vague beliefs to well-articulated theories. It would be useful to specify more precisely whether each of these is equally strong evidence against associationist views of conceptual development.

Like Gelman and Koenig, Gopnik and Nazzi (chap. 12) argue that perceptually based accounts of categorization are insufficient. While Gelman and Koenig point to ontology, function, intention, and other nonobvious properties as evidence for young children's theory-based categories, Gopnik and Nazzi focus mainly on causality, arguing that essentialism reflects a belief that there is some common cause for the characteristic properties of objects (see Strevens, 2000, for an argument that children (and adults) may hold nonessentialist causal beliefs). On their view, all of the theory-based evidence reflects a belief in some causal mechanism. In the case of biological natural kinds, for example, internal structure will be postulated as the causal mechanism; in the case of artifacts, the purpose for which an object is designed will be the cause. Causal mechanisms predict and explain correlations among features, and provide a reason for believing in the inductive potential of categories. Gopnik and Nazzi propose not only that an object's underlying cause predicts the correlations among its properties but that children can reason from patterns of correlations to cause. Rather than simply attending to patterns of correlations of features as a summary of object characteristics, these patterns can actually be used to infer causal relations. In marked contrast to Smith et al. (chap. 11), for Gopnik and Nazzi "perceptual features are important only insofar as they are an indicator of underlying causal powers" (p. 305). In this analysis, Gopnik and Nazzi provide an original theoretical framework and powerful models of children's reasoning about cause-and-effect relations.

Based on this analysis, one might expect their research to focus on ways that children's essentialist theories account for the correlated features of objects, but Gopnik and Nazzi instead provide an analysis of causal relations between one object and others. It is important to note that there are likely to be major differences between explanations of why intrinsic properties cohere and explanations of how one object exerts an effect on others. As Abelson (2001) has argued, there are significant differences in children's ability to reason about intrinsic versus extrinsic properties of objects. Gopnik and Nazzi provide a compelling model of extrinsic cause-and-effect relations, but they have yet to apply it to the problem of why an object has the features it does. That said, we turn to summarize the detailed and subtle argument they make.

A good analog for the kind of problem Gopnik and Nazzi consider is magnetism. Two similar objects, say two spoons, can differ in terms of which is attracted to a magnet, while two dissimilar objects can both be attracted. Gopnik and Nazzi argue that children set up "causal maps" of the world that go beyond whatever fundamental causal notions or perception may be hardwired. These learned causal maps consist of "abstract, coherent, representation[s] of the causal relationships among kinds of objects and events" (p. 307), where causal relationships between events are systematically related to patterns of conditional probabilities among those events. The model also takes into account whether an outcome was produced by some intervention apart from the causal relations being analyzed.

To test these ideas, Gopnik and colleagues have predicted that children categorize and name objects on the basis of perceived common causal powers. They describe a number of studies designed to test these ideas with children ranging in age from 18 months to 4 years. Many of the studies used a procedure involving a "blicket detector": An object the experimenter called a blicket was placed on a machine and made the machine light up and play music. The key question was what other objects children would consider to be blickets—objects that were perceptually similar but failed to set the machine off or less similar objects that made the machine react. Three- and 4-year-olds were equally likely to choose causally or perceptually similar objects. Two-year-olds selected perceptually similar objects more often but still chose
causally similar ones more often than a distractor object. In a control experiment, the contingencies were the same, but the experimenter held the object he called a blicket over the machine while he obviously pressed the machine with his other hand, apparently activating its lights and music. Now children routinely selected other blickets based on perceptual similarity. By analyzing the hand as an intervention that activated the machine, the same statistical association between the objects and the machine's activity was reinterpreted by the children. These initial studies established the basic findings that children name objects on the basis of their causal powers and that the causal maps they use go beyond simple statistical regularities.

As with any seminal work, these studies raise a number of questions. First, it is important to clarify the difference between the kind of analyses required to establish causal maps and the kind of statistical analyses assumed in the models of Smith et al. (chap. 11) and Mareschal (chap. 14): What is it about the Bayes nets, which work by establishing relations between conditional probabilities, that defies the assumptions of the associationist models? Where would the simpler statistics fail in terms of learning patterns of relations among conditional probabilities?

A second question concerns the domain specificity of the mechanism: What leads children to establish causal maps? Does a child need to suspect a causal relation in order to notice, represent, and analyze the relevant relations? If so, what would cause a child to initiate this analysis? To return to the point we made earlier about intrinsic versus extrinsic relations, seeing that something appears to make a machine light up might cause children to wonder what is happening or how it works. But simply observing a property of an object, for example, that a shirt is red or that a tiger has stripes, might not readily elicit the same desire for an explanation. In other words, seeing what appears to be an effect of one object on another might be an impetus to construct a causal analysis, while simply observing an object with its intrinsic properties might not. If a trigger is needed, then unless something draws their attention to a causal question, children might not routinely set up causal maps to explain why objects have the properties they do. If a trigger is not necessary, then is this propensity to set up causal maps totally unconstrained? Does a causal analysis run indiscriminately, continuously setting up relations between conditional probabilities for a huge number of coincidental events, with only a small percentage of them paying off? It seems more likely that some prescreening sets this analysis in motion—some sense that a causal question is of interest.

A third question, related to domain specificity, concerns the substantive content, such as core concepts and causal principles, of the theories. It is worth taking lay physics as a model of a theory and contrasting it with the causal theory of Gopnik and Nazzi argue children create to explain the blicket detector. In lay physics, the theory has content with substantive concepts and causal principles at issue, not just an abstract map of patterns of conditional probabilities: objects move about in space and time, two objects cannot occupy the same space at the same time, and so on. It is unlikely that children have any such principles to distinguish blickets from nonblickets. Another question this research raises then is how causal maps might provide the scaffolding for the more substantive concepts that a causal theory will ultimately have.

Naming and the Scope of Lexical Categories

We turn now to directly address questions about the role of language in category development. Waxman (chap. 9) tackles a fundamental question at the heart of word learning. Her chapter is an attempt to reconcile two competing theories of how children learn words: one posits a priori expectations that link each type of word with a particular type of meaning, and the other suggests that these expectations are not a priori but develop during the course of word learning. Waxman’s thoughtful solution to this debate is to suggest a combination of the two. She argues that children begin the process of word learning with an expectation, namely that words highlight commonalities between objects or, as she puts it, words serve as “invitations to form categories.” Experience, however, provides children with the necessary input to learn that certain grammatical form classes highlight certain categories. For example, nouns highlight object-based categories, and adjectives highlight property-based ones. Because languages mark grammatical form classes in different ways, children must learn the relevant linguistic conventions in their language before they will have expectations about specific word-category links.

Using a procedure that combines elements of a novelty-preference task with elements of a word-extension task, Waxman and her colleagues have demonstrated just such a progression from a general expectation about words to more refined expectations about particular grammatical form classes. They have shown that 11-month-olds have a general expectation that words highlight categories but do not yet distinguish between nouns and adjectives. Fourteen-month-olds, on the other hand, are a bit more advanced: they are beginning to work out the word-category links, expecting nouns to extend to object categories, but they are not yet sure about adjectives, extending them to either object-based or property-based categories. Presumably, older infants would also expect nouns to highlight object-based categories, but they would expect adjectives to extend only to property-based ones. Even very young children, then, most of whom have barely begun producing their first words, do seem to have an expectation that words highlight commonalities between objects (see also discussions in Markman, 1992; Markman & Hutchinson, 1984; Hollich, Hirsh-Pasek, & Golinkoff, 2000). It remains an open question how these expectations are formed.

Mervis, Pani, and Pani (chap. 10) also argue that words highlight commonalities between objects. Like Waxman, they argue that a child’s expectation about which features of a category are relevant to that category may differ from an adult’s. However, whereas Waxman focused on how inexperience with grammatical form class can lead a child to make errors in extension, Mervis et al. focus on how inexperience with adult-relevant features of an object can lead to extension errors. On their argument, a child may think that certain features of an object are critical to its identity, when in fact these features may be irrelevant from an adult perspective. For example, a child may overextend ball to a ball-shaped candle because, from the child’s point of view, its shape or rollability is its most important feature. Similarly, a child may fail to extend ball to a football because it lacks the shape associated with balls. This is not to say that children do not notice certain similarities between, for ex-
ample, the football and more typically shaped balls; they merely emphasize certain features they believe to be more relevant (in the sense of L. Bloom, 1993).

Mervis et al. offer a unique approach to studying this issue, with exceptionally detailed diary data from one child's acquisition of words in the bird domain. Concentrating on a single child's acquisition of words in a single hierarchical category has the advantage of allowing for the careful study of how input affects uptake and how new words change a category's structure. Not many children younger than 2 years old have the elaborated hierarchies of birds that Ari had, and they may, in fact, lack any equally elaborated hierarchy. Nonetheless, the principles by which he learned new words and incorporated them into his existing lexicon have implications for theories about how children learn words and the power that labels come to have.

Given that children's categories as embodied in their lexical categories may not overlap perfectly with those of adults, the question taken up in Mervis et al.'s chapter is how they come to coincide. Mervis et al. argue that this process is the result of a cooperative effort between the child and his or her communicative partners. On Mervis et al.'s view, the input strategy used can influence the extent to which a child incorporates new information into an existing lexical entry. A strategy that highlights the distinguishing adult feature seems to be more effective at changing early lexical categories than a strategy where an adult simply provides a new label for a misclassified object: highlighting the adult-relevant feature can make it clear why a new category assignment is appropriate for that object, thereby providing a hook for the child to begin forming a new category distinct from the known one. Later, children may learn a new label for a misclassified object without any supporting information, but this requires that the child have what Mervis et al. call the authority principle, or the understanding/belief that "there exist people (authorities) who know more about forms, functions, and form-function correlations than you do" (p. 248). Note that an "authority principle" could be invoked to explain children's acceptance of adult input in domains other than language. Such an assumption might facilitate the acquisition of knowledge much more generally. In terms of category learning, the authority principle allows children to accept a category assignment even when they do not understand the basis for it.

On our view, a willingness to accept a category label for objects that children would not have classified together on their own is crucial to conceptual development. No longer is the child bound by perceptual categories, or even by functional (Kemler Nelson et al., 2000) or causal (Gopnik & Sobel, 2000) categories formed on the basis of the child's own experience. Instead, children can form categories on the basis of language, with a simple category label serving as a placeholder for some possibly as-yet-unspecified similarity or cluster of similarities that makes members of that category coherent.

Putnam (1975) has suggested that a key element of adult communication is that we frequently refer to things without necessarily knowing the criteria for use of the terms we use. He writes that "every linguistic community . . . possesses at least some terms whose associated 'criteria' are known only to a subset of the speakers who acquire the terms, and whose use by the other speakers depends upon a structured cooperation between them and the speakers in the relevant subsets" (p. 126). In his words, language use requires a "division of linguistic labor." For example, although most adult speakers of English could identify most instances of the term tiger, few would be able to provide a compelling explanation or test for what makes something a tiger; for that, we rely on experts (although, in fact, there may be no scientifically valid essentialist explanation; Gelman & Medin, 1993). By analogy, children who accept that two things they would not have spontaneously categorized together are members of the same kind will expect those things to have unobservable properties in common—properties relating to ontology, causation, function, essentialism, and so on that are the fundamental determinants of what something is, what it does, what you can expect from it, and so on. Note that, as our earlier discussion of MacCarron and Callanan (1995) showed, children do not blindly make the decision to treat two things called by the same name as the same kind of thing. They use all of the available information in making their decision, and which information is most relevant will depend on the type of theories children have.

Mervis et al. argue that Ari did not accept new category labels for miscategorized objects until about 20 months, which could be interpreted as meaning that prior to 20 months, he did not expect categories to have nonobvious similarities or did not yet possess the authority principle. We think this is a conservative estimate. In a recent study by Graham, Kilbreath, and Welder (2001), for example, 13-month-olds were quite likely to use a novel object to imitate an action that produced an interesting sound if that object looked highly similar to a demonstration object that had earlier been shown to produce the sound (see also Baldwin, Markman, & Metzler, 1993). They were unlikely to attempt to perform the action if the object was low in perceptual similarity. However, when 13-month-olds heard the object low in perceptual similarity called by the same name as the demonstration object, they were just as likely to attempt to perform the target action. In other words, these very young infants—just beginning to produce their first words—already expect category members to share unforeseen similarities, and they can use language to categorize objects in nonobvious ways. Similarly, Waxman and Markow (1995) have shown that a novel label applied to members of the same superordinate category can facilitate the formation of a superordinate category in 12-month-olds.

The child's version of the division of linguistic labor seems more appropriately characterized as a continuum rather than an all-or-none ability, with Waxman's "invitation to form categories" toward one end and Mervis et al.'s authority principle toward the other. That is, very young infants may be able to form a new category on the basis of familiar objects sharing the same novel label, as in Waxman and Markow (1995), or novel objects sharing the same novel label, as in Graham et al. (2001). But they may not be able to form a new category for an initially misclassified object on the basis of linguistic information alone, as in the examples in Mervis et al. (chap. 10). Recategorizing an initially misclassified object into a known category on the basis of linguistic information may be an even more advanced ability, requiring the child not only to accept that nonobvious features bind two dissimilar things together but also to discount the perceptual features of one of the objects entirely.

We have recently begun a series of studies that may shed light on this issue (Jaswal & Markman, 2002). Briefly, infants are shown, for example, that a dog plays with a stick and that a cat plays with a ball of yarn. They are then shown a computer-
generated animal that adults and preschoolers have indicated looks very much like a dog (but that has some features of a cat), and infants are asked to show what it plays with. Some children hear this chimera called this cat, and some hear it referred to as this one. Twenty-four-month-olds who hear it referred to as this cat are quite likely to infer that it plays with yarn, while those who hear it referred to as this one are more likely to infer that it plays with a stick. For a category label to have this effect, children must be willing to set aside their own compelling, perceptually based classification in favor of a classification they do not understand. That is, they must accept "on faith" that there is a reason why an animal that looks very much like a dog is being called a cat. Whether children younger than 24 months can also use linguistic information in this way is the focus of ongoing work.

Of course, under most circumstances, using an object's appearance to determine its category membership is reliable (Landau, Smith, & Jones, 1988). A related objection is that in the real world, there are many more cues to an object's identity than the single, often purposely degraded or ambiguous image provided in many studies of category induction—cues like the way it moves, its smell, views from multiple vantage points, and so on (Jones & Smith, 1993). All of this is obviously true: there are very strong correspondences between appearance and identity; as adults, we may even think that there is a nonarbitrary relationship between the two (Medin & Ortony, 1989). Indeed, in a typical 2-year-old's productive vocabulary, most category labels represent categories whose members cohere perceptually (Samuelson & Smith, 1999). Ironically, we find that these objections actually bolster our case: despite the fact that appearances usually are a good clue as to category membership, and despite the fact that most count nouns encode perceptually similar objects, even very young children expect category members to share deeper similarities (though they may not know what those are), and they can use language to categorize objects in these nonobvious (and often underspecified) ways.

Conclusions: Prerequisites of a Theory of Conceptual Development

These six chapters, along with other recent advances in the field, point toward a set of principles, abilities, and phenomena that a complete theory of conceptual development must address. Extrapolating from this and other work, we propose that a coherent theory of children's categorization requires an integration of the following abilities of infants and young children.

Associative and Statistical Mechanisms

There are now several striking demonstrations of the power of associative mechanisms to pull out the correlational structure of objects in the environment (chaps. 4, 7, 11, 14), and even some work documenting infants' abilities to compute conditional probabilities over some kinds of events (Saffran, Aslin, & Newport, 1996). At a minimum, these computations of the relations between features provide useful partitions of the world. Are children ever limited to carving up the world into only those categories that cohere perceptually in this way? In our view, this remains an open question, but if there is such a limitation, it is only in early infancy. Moreover, there is even evidence that 10-month-olds don't or can't use such featural information for tracking numerical identity of objects and that they rely instead on spatial-temporal information (Xu & Carey, 1996). If infants rely solely on perceptual features to identify objects, why, then, would they resort to principles of lay physics for tracking object identity?

We agree that, throughout life, associative mechanisms supply crucial input into a system that constructs categories and that they provide information that is routinely used for identifying category members. What something looks like, sounds like, and feels like is important information about what it is. Important, but not definitive. We also agree that some of the mapping of categories to a given lexical item or to a given grammatical form class is accomplished in part by associative mechanisms. But existing models have a serious shortcoming: they treat words as features that enter into a correlational matrix. In fact, words are not features of objects; words are symbols that refer.

Causal/Theory Based Reasoning

Even infants encode at least some events in terms of causal relations, and there are several ways in which causality or theory-based reasoning play important roles in children's construction of categories and reasoning about kinds. We emphasize three points here.

1. Relevant ontological/theoretical/causal knowledge is recruited where possible and where needed, even by infants. A given category can be theory-based without requiring new causal principles to be developed for each category or categorization decision. Instead, relevant principles that have already been established can be recruited as appropriate for a given category (see chap. 7). Even young infants honor ontological distinctions such as animate/inanimate or object/substance. These core ontological principles influence their categorization and can even cause them to categorize in ways that contradict an analysis based solely on perceptual features.

2. Children believe in determinism. Even very young children probably have at least a vague sense that all things must have a cause. This is reflected in children's essentialist beliefs about categories. This belief can serve as a motivation to seek a deeper, nonperceptual explanation for why things have the obvious and nonobvious properties they do. It also makes it plausible for children that two things that don't seem much alike could still be members of the same category. That is, children are willing to defy the output of the associative analysis on perceptual features because they take it on faith that there are deeper causes for things to be the way they are.

3. There are some potentially domain specific learning mechanisms for constructing theory-based knowledge. One form that these mechanisms take might be tied to specific foundational theories, such as lay physics or lay psychology, that infants may hold. A second form is the one that Gopnik and Nazzl advocate—a computational ability that enables the construction of causal relations.
The Division of Linguistic Labor

Even adults accept and use terms without being able to provide the scientific criteria for their use (e.g., Could you perform a test to distinguish between gold and fool’s gold?), accepting what Putnam (1975) has called the division of linguistic labor. We all benefit from the accumulated knowledge of our culture, handed down to us, in part, through the categories that have proved useful to others in the past. The labels that objects are given can inform us about those categories even when we may not understand their basis. Thus, as many have argued, language serves as a mechanism for the cultural transmission of knowledge (Gelman, Hollander, Star, & Heyman, 2000; Gentner & Boroditsky, 2000; Tomasello, 1999).

Children with very little linguistic experience can use language to carve the world up differently from how they would do so spontaneously, suggesting that they also adhere to something like a division of linguistic labor. On our view, the division of linguistic labor and essentialism go hand in hand, complementing each other and allowing for the formation of categories that are not perceptually or even experientially based. That is, essentialist beliefs lead children to expect that category members share deeper, nonobvious similarities. A simple category label can serve as a placeholder for a theory about why two things are members of the same kind and can be used to make inferences when kind is relevant to the inference. But for language to have this power, children must sometimes be willing to accept counterintuitive and baffling labels that defy the output generated by lower-level associative mechanisms (e.g., Why is an eel a fish? Why are both tables and lamps furniture?). That children regularly do so suggests that they accept a division of linguistic labor and can use labels to learn categories that fly in the face of perceptual similarity.

In sum, the ability to compute statistics that partition the world into objects with correlated features is an important source of input into any system that must identify objects. But the ability of those domain-general learning mechanisms to account for apparently uniquely human assumptions about essentialism, causality, and nonobvious features—assumptions that enable deeper thinking and reasoning about objects—is questionable. It is these assumptions, in combination with the division of linguistic labor, that allow children to form categories they could not form on their own and to benefit from the accumulated knowledge of their culture.

Acknowledgments

We thank Susan Johnson and Josh Tenenbaum for helpful discussions about issues raised in this chapter.

References


Early Category and Concept Development
Making Sense of the Blooming, Buzzing Confusion

Edited by
David H. Rakison and Lisa M. Oakes

OXFORD UNIVERSITY PRESS
2003