5. Language and categorization: The acquisition of natural kind terms

SUSAN A. GELMAN AND JOHN D. COLEY

What is striking about human categories is their diversity. They range from the simplest classification of a face or color to the most carefully constructed taxonomic grouping. Considering this diversity, many are struck by the apparent gap between the simple, intuitive categories formed by children and the complex, theory-laden categories of educated adults (e.g., Inhelder & Piaget, 1964; Quine, 1969; Vygotsky, 1962).

In this chapter we first argue that despite a number of salient differences between children’s categories and those of adults, there are important parallels between the two. Both are informed by an ability to overlook salient appearances, an attention to nonobvious properties, and the potential to draw many new inferences about the unknown. Both the initial groupings of the prescientific child and the most thoughtful, theory-laden classifications of the adult extend knowledge in important ways. To use Quine’s terminology, both children and adults form “theoretical kinds.” Second, we address the role of language in the formation of theoretical kinds. Although the structure of everyday object categories (e.g., dog, hammer, oak tree, and computer) is traditionally thought to result from the structure of the world and/or the nature of human perception and cognition, we will present evidence that language is also critical and that how objects are named helps determine the structure of the categories they fall into.

The chapter has three sections. In the first, we set forth our as-

sumptions about the nature of categories for adults. We review recent analyses suggesting that categories are enriched and informed by intuitive theories, summarizing arguments from psychology, philosophy, and linguistics to converge on the point that human categories extend far beyond observable similarities. In the second section, we demonstrate that this analysis of adults can be extended fruitfully to children. Contrary to most standard views of children, and despite children’s relatively modest scientific knowledge, many of their categories have theory-laden properties. In the final section, we examine the role of language in the formation and structure of children’s theory-laden categories.

Natural kinds as theory-laden categories

One critical feature of human categorization is that in forming categories we strive to incorporate our theories about, and explanations of, the world (Quine, 1969). Quine distinguishes between intuitive kinds (e.g., groupings based on color), which even rats form with ease, and theoretical kinds (e.g., the classification of bats as mammals rather than as birds), which require a restructuring of our innate sense of similarity. Theoretical kinds result from our attempts to understand the environment, not simply to organize and catalog it. Such categories hang together to the extent that they can explain and predict phenomena, instead of simply describing them. Mill (1843) aptly explained this function as follows: “Classification ... is a contrivance for the best possible ordering of the ideas of objects in our minds; ... in such a way as shall give us the greatest command over our knowledge already acquired, and lead most directly to the acquisition of more” (p. 432). In the rest of this section, we provide recent arguments and evidence that theoretical kinds are indeed central to much of adult thought.

Theory-laden categories

Arguments for the importance of theories in adult categories begin with the observation that similarity alone cannot explain why things belong to a given category. In order to calculate similarity, it is necessary to have a theory of which features are important and how they should be weighted. Murphy and Medin (1985) present two arguments for the inadequacy of similarity for categorization (see
also Goodman, 1972). First, the similarity of any two objects depends on which properties of the objects are under consideration. As they point out, "Any two entities can be arbitrarily similar or dissimilar by changing the criterion of what counts as a relevant attribute" (p. 292). Two seemingly disparate objects, such as a blender and a sparrow, could be considered highly similar if certain attributes are highlighted. For example, both are less than 10,000 years old, both weigh less than 2,000 kilograms, both can be found in North America, both visible without magnification, and so on. Second, even if the set of possible attributes is somehow constrained, similarity judgments depend on the relative importance of those attributes. There is no theory-neutral weighting of properties. For example, both habitat and method of breathing may be important attributes for classifying species. If habitat is given more weight, humpback whales might be classified with fish. But if breathing method is considered more important, humpback whales would be considered mammals. Similarity alone is insufficient to hold categories together because it does not by itself provide adequate constraints on category membership.

Rips (1989) provides empirical evidence for the distinction between similarity judgments and categorization, showing that judgments of similarity and category membership vary independently, and therefore similarity cannot completely account for categorization. In one experiment, adult subjects said that an unknown object three inches in diameter1 was more similar to a quarter than a pizza, but more likely to be a pizza than a quarter. Presumably, theoretical beliefs concerning the range of possible sizes for pizzas and quarters determined the subjects' category judgments. In another experiment, subjects read stories about animals undergoing accidental transformations (e.g., following exposure to toxic chemicals). For example, the birdlike "sorp" (eats seeds and berries, lives in a nest, has two wings, two legs, feathers, etc.) was rated by a control group as much more similar to a bird than to an insect and much more likely to be a bird. After exposure to the chemicals, the sorp was transformed into an insectlike creature (it began to eat only the nectar of flowers, took shelter by sticking to the underside of leaves, sprouted new wings made of transparent membranes, developed two more pairs of legs and a brittle shell, etc.), but it was still able to mate and produce normal sorp offspring. After reading this tale, the subjects rated the transformed sorp as more similar to an insect than a bird, but still more likely to be a bird, despite the changes. On the basis of these and other experiments, Rips argues that the independence of the categorization and similarity responses shows that one cannot be reduced to the other.

The claim, then, is that category membership is dictated not by perceptual similarity but by deep or predictive properties. Categories are psychologically coherent not because of surface similarity among members, but because of explanatory theories that underlie them. "Concepts are viewed as embedded in theories and are coherent to the extent that they fit people's background knowledge or naive theories about the world" (Medin & Wattenmaker, 1987, p. 58). So, to return to our previous examples, whales are mammals because they share properties with other mammals that modern biology deems important, such as type of metabolism and method of reproduction. Conversely, sparrows and blenders do not form a coherent category because we are hard pressed to come up with an explanatory theory linking the two. Note that theory here refers to "any of a host of 'mental explanations' [including scripts, rules, and causal explanations], rather than a complete, organized, scientific account" (Murphy & Medin, 1985, p. 290).²

The theory view is perhaps best understood when contrasted with other possible positions. Some have argued a nominalist position, that human categories represent an essentially arbitrary framework imposed on the world. This implies that the psychological act of categorization is one of invention, as described by Bruner, Goodnow, and Austin (1956):

Science and common-sense inquiry alike do not discover the ways in which events are grouped in the world; they invent ways of grouping. . . . Do such categories as tomatoes, lions, snobs, atoms, and mammalia exist? In so far as they have been invented and found applicable to instances of nature, they do. They exist as inventions, not as discoveries. (p. 7)

Nominalists argue that every object has many more discriminable attributes than we could take into account in any categorization and that therefore we must select which attributes to consider. Because attention to different attributes could lead to different categories, categorization is essentially an inventive act.

Others have argued a realist position, that categories are direct reflections of a structured world. This perspective is typified by the work of Rosch and Mervis (1975), who contend that "division of the
world into categories is not arbitrary. The basic category cuts in the world are those which separate the information-rich bundles of attributes which form natural discontinuities" (p. 602). Realists hold that objects naturally fall into groups, and in such groups members are similar to one another and dissimilar to members of other groups. On this account, the attributes important for categorization exist in the environment as correlated bundles (for reviews see Mervis & Rosch, 1981; Rosch, 1978).

The nominalist position stresses the arbitrary nature of categories; the realist position places heavy emphasis on perceptual features devoid of context or explanation. The theory-laden position supplants this controversy, making use of both the notion of inven-
tiveness in categorization and the notion of constraints inherent in the environment. For example, the theory-based grouping of bats as mammals is partly an inventive decision (because they could be classified differently), yet follows from our attempts to capture regularities in the environment (e.g., deeper qualities, such as means of bearing young). Our categories must be consistent with patterns in the environment, yet are not limited to perceptually salient features. Perceptual cues often signal the presence of these deeper qualities, but are not themselves the basis of category structure.

Properties of natural kind categories

Taking a broad view, theories are important for nearly all human categories (Murphy & Medin, 1985). For example, even an ad hoc category such as things to take out of a burning house (including children, jewelry, paintings, and book manuscripts; see Barsalou, 1983) gains coherence by means of a theory or goal that tells us what people value. Without considering the underlying goal that holds these items together, they have no similarities to speak of. But when the goal that relates category members is explicitly stated, the category becomes coherent. Subjects can even judge the typicality of category instances (as in more traditional categories, such as birds).

Nonetheless, on an intuitive level it would seem that categories vary in their theory-ladenness. Although many categories have psychological coherence, only some categories attempt to carve nature at its joints. Although many categories depend on theories, only a subset promotes an inexhaustible set of nonobvious inferences and warrants scientific investigation. It is useful here to invoke a dis-
tinction between natural kinds (e.g., tiger, lemon, gold, electron) and other categories (e.g., chair, box, things to take out of a burning house) (Locke, 1706/1965). What are the properties that distinguish natural kinds from other, less theory-laden categories? Below we present six properties that characterize natural kinds: rich inductive potential, nonobvious basis, essence, existence of anomalies, division of linguistic labor, and corrigibility. These interrelated properties convey the psychological nature of natural kinds. They are derived from discussions of natural kinds in philosophy and linguistics by Kripke (1971, 1972), Mill (1843), Putnam (1970, 1973), Quine (1969), Schwartz (1977, 1979), and others. Our aim is to detail with some precision what natural kind concepts are like in ordinary language.

Rich inductive potential. People construct categories with the implicit (perhaps unconscious) goal of learning as much as possible about the objects being classified (Jevons, 1877; Mill, 1843; Ruse, 1969). For example, if we learn that X is a "cat," we infer that it has many important properties in common with other cats, including diet, body temperature, genetic structure, and internal organs. We can even induce previously unknown properties. For example, if we discover that one cat has a substance called "cytosine" inside, we may then decide that other cats also contain this substance. Theory-laden cate-
gories encourage us to discover more similarities among the objects being classified than we would ever discover otherwise.

Indirect evidence for the importance of inductive potential can be found by examining which kinds of groupings we prefer over others. For example, we classify animals by shape and covering because these properties reliably predict other, more central properties such as habitat, diet, and means of locomotion. We would probably not even think to classify animals solely by color (e.g., considering brown cats, brown bears, and brown birds to be members of the same species, but white cats, white bears, and white birds to be members of another species), because color is not predictive of many other properties for animals.

Nonobvious basis. Natural kind categories promote inductive infer-
ences in part because they capture deep similarities that are not al-
ways immediately obvious (see Wellman & Gelman, 1988). As we have seen, theories dictate which similarities count as important. One striking example is that of marsupial mice – small, ratlike an-
imals that are classified as kangaroos because they have pouches and lack placentas (Quine, 1969). Marsupial mice are classified on the basis of underlying properties; they are not classified according to their most obvious surface features.

The importance of nonobvious features holds not only for unusual instances (such as marsupial mice), but also for more everyday entities. For example, we expect all tigers to have a particular genetic structure, chemical makeup, and bone structure in common, not just stripes; we expect all water to have a certain chemical structure in common, not just potability and transparency. Nonobvious features are also critical for understanding two other ubiquitous phenomena: developmental change (note the difference between a caterpillar and a butterfly, an ugly duckling and a beautiful swan, or an infant and adult human) and sexual dimorphism (males and females of many species differ radically in size, outer coloring, and/or morphology). Common nonobvious features lead us to classify caterpillars and butterflies together, or peacocks and peahens together, despite contrasting outward appearances.

**Essence.** Medin (1989; Medin & Ortony, 1989) has suggested that adults believe that certain categories have an essence, or a unique underlying property that is responsible for other similarities that category members share (Goosens, 1977). He calls this belief "psychological essentialism." In contrast to essentialism, which is a metaphysical claim about the structure of the world, psychological essentialism is a claim about human cognition. For example, although tigers share many properties (a certain DNA structure, bone structure, chemical makeup, appearance, and behavior), people may believe that DNA is the essence of a tiger because it causes the other properties. We do not even need to know what the essence is to believe it exists (see the subsection on the division of linguistic labor), although we presume that science will discover the essence eventually.

Although the notion of an essence is objectionable from an evolutionary perspective (because having an essence implies that the essence—and so the category—is immutable; Mayr, 1988), it has appeared repeatedly throughout recorded history (Kelly, 1989). Nonscientific cultures also appear to treat categories as embodying an essence. Using folk taxonomies as evidence, Atran (1987a, p. 197) suggests that there is "a (universal) presumption to the effect that visible organic types have underlying natures and that these essences provide a principle of natural causality for manifest organic regularity. . . . Cross-culturally, people presume that living kinds have (possibly unknown) natures with propensities responsible for the readily perceived regularities of those kinds."

**Existence of anomalies.** Because nonobvious features are critical to category membership, natural kind categories include anomalies, or members that differ greatly from the category prototype and may outwardly seem not to belong (Putnam, 1970; Quine, 1969). For example, we readily accept that birds include ostriches and penguins, or that fish include eels and flounder. Some of the properties that are usually used to identify category members can be missing, but we will still agree that the object is a member of the category if there is reason to believe that the more essential, more explanatory properties still hold. Albino dwarf tigers might lack many of the properties we usually associate with tigers: They would not have stripes, they would not be ferocious, they would probably look and act more like domestic cats than tigers. Because hidden properties are more important for natural kinds, surface similarity is only a fallible guide to category membership.

**Division of linguistic labor.** The average speaker recognizes that theory-laden knowledge is critical for the appropriate use of natural kind terms, even when that same speaker does not possess the underlying theoretical knowledge. Thus, one does not have to be an expert to treat categories as natural kinds. This point is critical, because it suggests that theories are considered integral to natural kinds and are not simply an accidental by-product of schooling or expertise. Putnam (1973) explains this as follows:

The features that are generally thought to be present in connection with a general name—necessary and sufficient conditions for membership in the extension, ways of recognizing whether something is in the extension, etc.—are all present in the linguistic community considered as a collective body; but that collective body divides the "labor" of knowing and employing these various parts of the "meaning" of [e.g.,] "gold." (p. 125)

Thus, the basis of a category label can go far beyond what is known by most users of the word. For example, many of us do not know the difference between elms and beeches, but a botanist does. Even
though we may not know what makes these trees different, we assume there is an underlying distinction that an expert would know. Putnam makes the analogy with tools requiring cooperative activity for their use (e.g., steamships) as contrasted with tools that can be used by an individual (e.g., hammers). He suggests that many words function like cooperative tools.

_Corrigibility_. Natural kind categories are always open to revision (Putnam, 1970, 1973; Quine, 1969). As long as theories are open to change, categories too will be open to change. For example, whales used to be considered fish, but as people learned more about them, they reclassified them as mammals. The category changed because people sought a grouping that would reflect deeper properties such as body temperature and breathing patterns. In one sense all categories are corrigeable in that they undergo historical change. To give a recent example, in the NCAA basketball tournament, shots from beyond 18 feet, 8 inches were classified as “two-point shots” before the 1987 tournament and as “three-point-shots” thereafter. However, in a stricter sense corrigibility implies that classifications were previously either in error or incomplete. In this sense, the categories of basketball shots are not corrigible: It is not true that shots beyond 18 feet, 8 inches were actually three-point shots before 1987 and that the NCAA classification was wrong. In contrast, whales are and always have been mammals (at least according to current theories), and previous classifications of whales as fish were incorrect.

Which categories are natural kinds?

The epitome of natural kinds are scientific concepts, which by definition have a persistent commitment to go beyond similarity (see also Kelly, 1989). For example, in the modern conception of species, the mechanisms underlying the observable similarities and differences across species are primary (Ghiselin, 1969). As a consequence, surface similarities at times become irrelevant, taking a back seat to the underlying mechanisms, such as genealogy, reproductive isolation, or chromosomal structure. Ghiselin characterizes modern conceptions of species as follows: “Instead of finding patterns in nature and deciding that because of their conspicuousness they seem important, we discover the underlying mechanisms that impose order on natural phenomena, . . . then derive the structure of our classification systems from this understanding” (1969, p. 83). As this example illustrates, the modern species concept in biology is notable for its search for inductive potential, appeal to nonobvious properties, incorporation of anomalies, and so forth. These are qualities that are not evident in ad hoc categories such as things to take out of a burning house.

We propose that there are striking parallels between scientific natural kind categories (such as biological species concepts) and a subset of ordinary categories. In suggesting this analogy, we distinguish category content from category function. Clearly, content differences exist between scientific and ordinary classifications (Dupré, 1981; Morris, 1979). Ordinary people lack extensive scientific knowledge, so many words in ordinary language fail to map precisely onto scientific categories. For example, pterodactyls are not technically dinosaurs, but they are included in a box of “Dinosaur Honey Graham Cookies” sold commercially and in a collection of “dinosaur magnets” at an Ann Arbor store selling kitchen accessories. Similarly, the lay distinction between fruit and vegetables has no biological correlate. It should not be surprising that the content of ordinary language categories at times fails to mesh with the classifications of science, especially given that scientific categories are continually changing with the addition of new information.

However, the properties that make a category a natural kind concern how the category functions and do not determine which particular properties are critical for category membership. Despite differences in content, many categories function in similar ways for both scientists and nonscientists. Ghiselin (1969, pp. 87–88) notes that “the purpose of a [scientific] classification is not the accurate pigeonholing or identification of enzymes or dried specimens, but the assertion of meaningful propositions about laws of nature and particular events.” Or in Stephen Jay Gould’s words, “Classifications are theories about the basis of natural order, not dull catalogues compiled only to avoid chaos” (1989, p. 96). Similarly, in everyday life much – perhaps most – of what we learn is by inference (Holland, Holyoak, Nisbett, & Thagard, 1986), and categories provide an invaluable framework for such inferences. So although the content of the dinosaur category probably varies between expert and novice, the function remains the same, and most likely both novice and expert would expect deep, nonobvious similarities among dinosaurs.
Thus, natural kind categories are common in ordinary language. Many categories of naturally occurring objects or substances (e.g., fish, roses, gold) have all six properties, for ordinary adults. Even more abstract categories can have many natural kind properties. For many people intelligence is one such example: It has rich inductive potential (an intelligent person will probably perform well in school, have a successful career, etc.), a nonobvious basis (intelligence cannot be judged by appearances), perhaps an essence (such as a "g" factor), a division of linguistic labor (those who administer IQ tests are considered best equipped to determine whether someone is truly a genius), corrigibility (as evidenced by long-standing debates as to whether intelligence is just one factor or many factors), even anomalies (e.g., "idiot savants," who are geniuses of a sort yet seem not to be; learning-disabled children who may actually be very intelligent). Categories that have a full set of theory-laden properties can be considered natural kinds in that they aim to capture, as best as possible, the true structure of the world (see Jevons, 1877; Mill, 1843).

Many categories are assumed to be richly structured in the same way, even before all the evidence is in. For example, conceptions of certain disorders such as autism and dyslexia contain the hope that there is an underlying essence that, when discovered, will yield a rich array of inductive inferences concerning etiology, course of progression, and (ultimately) remedies. In other words, debates about how to define "autism" are not simply semantic, but rather concern which grouping will be most fruitful in discovering more about the problem.

Nonetheless, not all categories are theory-laden to the same degree. Categories that fall at the nontheory end of the continuum are sometimes referred to as artificial classifications, in that they are based on an arbitrary set of properties. They include categories defined by a single property, such as white things or objects bigger than a breadbox (see Mill, 1843). Ruse (1969) expresses this as follows: "An empty, arbitrary, or unreal concept yields nothing that was not already built into the definition" (p. 108). Because categories such as white things are structured around a single property with little predictive power, they fail to promote inductive inferences (e.g., if you learn something new about a white mouse, it is unlikely to generalize to a white fence), their basis is entirely obvious (resting solely on color), they have few if any anomalies (except for unusual viewing conditions, there are no white things that do not appear to be white), no division of linguistic labor (any intelligent, sighted person is qualified to determine what belongs in the category), no essence (there is no deeper property in common to all white things), and no corrigibility (we will never discover that light blue objects actually belong in the category of white things).

We propose that categories expressed in ordinary language vary along a continuum in the extent to which they can be considered natural kinds. Some categories (e.g., computers) have a nonobvious basis (it is its inner workings rather than outer appearance that makes something a computer), rich inductive potential (witness the emergence of the field of computer science), division of linguistic labor (a computer technician is better equipped to determine whether a new machine is a computer or some other device), even anomalies (such as cars with built-in computers), but probably do not have an underlying essence (since what makes something a computer changes over time). Other categories (e.g., descendents of George Washington) have a nonobvious basis (one cannot tell who belongs to the category by just looking), division of linguistic labor (a genealogist or other family members would be most accurate at identifying members of the category), and an essence (namely, a certain line of descent and genetic material), but little rich inductive potential or corrigibility. Still others (objects weighing 256 grams) may have a nonobvious basis, but few if any of the other theory-laden properties. And many common artifact categories (e.g., cup, chair) include anomalies (e.g., bowl-like cups, beanbag chairs) but none of the other theory-laden properties. In sum, there are a wide range of categories that have one or more theory-laden properties, including categories that are not traditionally thought of as embedded in theories. For example, the concept run batted in in baseball has many theory-laden properties, although in standard parlance it would be rare to say one has a "theory" of baseball (Siegler, 1989).

Knowing whether a category is a natural kind, an artificial classification, or something in-between has important consequences. The extent to which one considers a category to be theory-laden has direct ramifications for whether it is used to generate hypotheses and make inductive inferences. Cases of prejudice may arise when a rather limited distinction is assumed to be a natural kind distinction with an underlying essence. (For example, gender and race differences are heightened and exaggerated beyond their true basis, perhaps in conformity with the view that these are natural kind distinctions.)
Or in the case of medical disorders, if a disease is assumed to be
an artificial classification rather than a natural kind (e.g., a rash),
physicians will not bother to look for underlying commonalities.

The relation between theory and similarity

So far we have characterized "theory-based" natural kind categories
in contrast to "similarity-based" categories. We have argued that
human categories are not limited to perceptual similarity; they strive
to capture how the world works. The world is complex, so the
groupings we first notice do not always yield the deepest, most so-
plicated level of understanding, but rather undergo change with
increased knowledge and experience.

Nonetheless, perceptual similarity is often an important indicator
of theoretically relevant variables. In other words, similarity and
theory typically converge (see also Mayr, 1957; Medin, 1989). If our
immediate perceptions were not extremely good cues for locating
food, dangerous animals, or other members of our species, we would
not survive. Presumably we evolved such that the perceptual sim-
ilarities we find most salient are the ones most important for our
survival. As Quine put it, "Our innate standards of perceptual sim-
ilarity show a gratifying tendency to run with the grain of nature.
This concurrence is accountable, surely, to natural selection" (1974,
p. 19). Surface similarities may also lead us to look for deeper sim-
ilarities.

Moreover, when similarity and theory do not converge, each is
useful in its own way. The two are complementary sources of in-
formation. Similarity is most useful as a rough guide to identifying
what something is, whereas theory provides the underlying basis of
the classification. For example, birds are often identified by their
feather color, yet feather color is not the causal force behind the
classification. In the absence of more detailed knowledge, similarity
is a powerful default, even for adults (see Chi, Feltovich, & Glaser,
1981). This distinction between the underlying basis of a category
and reliable cues for identification emerges repeatedly in disuc-
sions of classification (see Atran, 1987b, conceptual identity vs.
conceptual access; Miller & Johnson-Laird, 1976, and Smith & Medin,
1981, conceptual core vs. identification procedure; Sokal, 1977, es-

dablishing a category vs. allocating objects to the category).

In short, it is important to realize that there are two separate is-
sues here. The first is: What information does one use to decide
what something is? In the absence of an explicit label, the answer
to that will nearly always be appearances – it may be uninformed
overall similarity, it might involve attention to a theoretically im-
portant subset of features (e.g., those indicating animacy; see Mas-
sey & Gelman, 1988), or it may even be similarity that requires ex-
tensive training, as in reading medical x-rays and classifying them
as x-rays of healthy versus x-rays of diseased tissue. The second
issue is: What are the structure and consequences of a category?
This is where an analysis of perceptual similarity is insufficient. Cat-
egories have implications that extend far beyond the features used
to identify members.

Summary

The distinction between theory-laden and non-theory-laden cate-
gegories has its roots in a distinction between natural and artificial
systems of classification (Knight, 1981; Markman, 1989). Mill (1843)
distinguishes between classifications that reflect patterns in nature
(natural kinds such as horse or animal) and those we construct for
our own convenience (artificial classes such as white things). Natural
types pick up rich clusters of information that extend far beyond
our original characterization; in contrast, artificial classes capture
simply the property on which they are based (e.g., for white things,
the property of being white). Artificial classifications name for the
sake of expressing a certain quality. They do not change with the
addition of new information, as do natural kinds. We discover nat-
ural classification systems; we construct artificial ones (Ghiselin, 1969).

A similar distinction has been proposed by Jevons (1877), who
remarks that "deep correlations, or in other terms deep uniformities
or laws of nature, will be disclosed by any well chosen and pro-
found system of classification" (p. 676). Whereas natural classification
systems aim to uncover a deeper reality, artificial classification
systems are useful because of the ease and convenience with which
they are put into practice. Although we strive for natural sys-
tems, we retain the artificial ones for our first attempts at classifying
and when we need something convenient that will work most (though
not all) of the time.

In this section we have proposed that natural kinds are theory-
laden categories exhibiting the following six properties: rich induc-
tive potential, nonobvious basis, underlying essence, anomalies, division of linguistic labor, and corrigibility. Categories vary along a continuum in the degree to which they embody these properties. The net implication is that a vast set of ordinary concepts requires attention to the nonobvious.

Children's understanding of natural kind categories

The view that many of adults' categories are theory-laden natural kinds raises a critical developmental question: How do children learn categories that have a nonperceptual basis, anomalous members, hidden essences, and/or the potential for unforeseen inductions? The standard developmental picture of young children might predict that they could not appreciate theory-laden, nonobvious concepts such as these. A vast array of research in cognitive development demonstrates that children attend to superficial appearances even when they are misleading (e.g., Bruner, Olver, & Greenfield, 1967; Flavell, 1963, 1977; Inhelder & Piaget, 1964; Lively & Bromley, 1973; Peevers & Secord, 1973; Piaget, 1970; Smiley & Brown, 1979). For example, in Piaget's conservation task, children see a volume of liquid being poured from a short container to a tall container. Nothing is added or taken away, but preschoolers insist that there is more liquid in the taller container because it appears to have more. Given children's attention to appearances, many have argued that an appreciation for theory-laden categories may emerge later in development. As Neisser (1987) puts it, "It is intriguing to discover that the two major alternatives considered by category theorists - featural similarities and theory-based definitions - correspond to points on a developmental continuum" (p. 6).

Another potential difficulty with early natural kind categories is that children have a relatively sparse understanding of science. In a detailed study of children's biological knowledge, Carey (1982, 1985) suggests that an appreciation for natural kinds might not develop until about age 10. She explains as follows:

First, the child may need a minimal understanding of the workings of science, in general, before he has the concept of a natural kind term. Second, the child (or adult) may need a minimal knowledge of a particular science before a natural kind term in the theory of that science is recognized as such. For instance, Putnam (1962) suggests that atom changed from a classical term meaning "smallest indivisible particle of matter" to its present natural kind status only when enough physics was known for atom actually to refer to atoms. Young children are markedly deficient in both reflections of scientific knowledge - the understanding of science in general and of particular theories as well. (1982, pp. 383-384)

Carey argues that knowledge about science in general and scientific domains in particular is necessary for a theoretically based conceptual structure to emerge and that without this requisite knowledge, children's categories cannot be characterized as natural kinds.

Nonetheless, the problems with similarity-based categories proposed for adults (Murphy & Medin, 1985) apply equally (if not more so) to children. Children must have some basis for constraining or weighting the features in the environment if they are to construct categories that make any sense. Moreover, it would be particularly useful for children to realize the rich inductive potential of categories from an early age, when they are first learning and organizing large amounts of knowledge.

We propose that children do appreciate natural kind categories well before the onset of formal schooling or the acquisition of extensive domain-specific knowledge. In particular, children use language to help set up categories that function like adult natural kind categories - as if they were theory-based - even though children have not yet filled in the particulars of the theory.

Past studies are limited in what they tell us about children's understanding of natural kind categories. First, past work assumed that categories do not substantially differ from one another. On this view the study of one sort of category is generalized to all the child's classification schemes and abilities. Categories that were studied previously often included wholly perceptual and/or arbitrary groupings, rather than natural kinds. For example, children might be asked to reason about blue circles and red squares (see Inhelder & Piaget, 1964). The class of blue circles is merely a grouping based on perceptual attributes; it does not capture nonobvious similarities, nor does it draw on richer sorts of knowledge children have about the world. However, as discussed earlier, categories differ in the richness of their structure and resemblance to natural kinds. To avoid this limitation, the categories to be studied should include nonobvious similarities and should tie into world knowledge (e.g., natural kinds).
Second, past work typically assumed that category knowledge can best be tapped by asking children to sort or classify objects. Tasks used previously most often required children to construct a classification (e.g., "Put together the things that go together") – thus emphasizing similarity – rather than to reason about or draw inferences from a category – thereby emphasizing the nonobvious. When we ask children simply which objects belong together, we are neglecting the deeper questions of whether the grouped objects form a motivated category and what the consequences are of having such a category. It seems clear that concepts comprise more than knowledge of which instances belong in a category (see, e.g., Salmon, 1981). Indeed, as argued earlier, inductions about completely new instances and properties reveal important aspects of category functioning.

In the remainder of this section we review developmental studies that avoided both of these difficulties. When the categories studied include natural kinds, and when children’s deeper understanding is probed, these studies show that young children (ages 2½ to 5 years) have an appreciation for categories that promote rich inductive inferences, share nonobvious properties that do not always correspond to outward appearances, have essential natures, and permit anomalous members. Although many of these features overlap, we separate them here for the sake of clarity. We will also consider whether for children, natural kind categories are subject to division of linguistic labor, and are subject to revision (exhibit corrigibility).

Rich inductive potential

Content versus function. Earlier we distinguished between content and function when comparing scientific and ordinary classification systems. Although scientific categories often differ from everyday categories in content, their functions are similar; both types of category serve to guide inferences beyond what is already known. Here we propose a similar relation between adults’ and children’s categories: Although children may rely on different bases for classification than adults (e.g., Inhelder & Piaget, 1964; Keil & Battersman, 1984; Smiley & Brown, 1979; Smith & Kemler, 1977; Vygotsky, 1962), their categories fulfill similar functions. Later we describe research demonstrating that for children as young as 2½, categories promote extraor-

dinarily rich inductive inferences and that this is true even in the absence of any perceptual support.

Categories and induction. As mentioned earlier, a major function of categories is to promote inferences beyond the obvious. There are two ways this happens. First, categories promote inferences about nonobvious properties, such as “produces insulin” or “has a spleen inside.” Such properties are not readily apparent from the outward appearance of an object. Second, even category members that do not look much alike share important properties (e.g., both a robin and an ostrich lay eggs, have warm blood, etc.).

To examine children’s understanding that categories promote inferences, Gelman and Markman (1986) devised a task designed to tap children’s inductive inferences. Induction entails making an inference beyond what one knows with certainty and contrasts with deduction, in which conclusions follow with certainty from the premises (Skyrms, 1975).

Gelman and Markman (1986) presented preschool children and adults with pictures depicting a wide range of natural kind categories, including snakes, dinosaurs, birds, squirrels, sand, gold, and diamonds, among others. Twenty sets of three pictures each were used. Each set was constructed so that the third picture closely resembled one of the first two pictures but was from the same category as the other. Thus, each item pitted perceptual similarity and category membership against one another. Of interest was whether children would draw inferences from one picture to another on the basis of outward appearance or natural kind category membership.

Two pictures from different natural kind categories (e.g., a colorful tropical fish and a gray dolphin) served as targets. These target pictures were named (in this case, fish and dolphin), and the subjects learned an unfamiliar nonobvious property about each. A broad range of important, nonobvious properties, concerning such things as behavior (“eats grass”), physical properties (“melts in an oven”), function (“has cold blood”), and origin (“comes from a mountain”), was used. In this example, the children were told, “This fish [i.e., tropical fish] stays underwater to breathe” and “This dolphin pops above the water to breathe.” (Pretesting confirmed that 4-year-olds did not previously know which property applied to each animal.) Children were then shown a test picture, in this case a gray shark. In all cases, the test picture belonged to the same category as one of the
targets (the shark and the tropical fish are both fish) but looked much more like the target picture from the other category (the gray shark and the gray dolphin look much more alike than the shark and the tropical fish). Adult ratings were collected to confirm perceptual similarity relations. After being shown the picture of the shark and told that it was a fish, the subjects were asked whether it stayed underwater to breathe like the fish or popped above the water to breathe like the dolphin.

When adults performed this task (with more difficult properties, such as “This bird’s heart has a right aortic arch only” for a flamingo), they based their inferences on category membership 86% of the time (i.e., they inferred that pictures from the same natural kind category shared novel nonobvious properties). As expected, this performance significantly exceeds what one would expect by chance. Furthermore, the subjects were highly confident of their answers, giving a mean confidence rating of 5.8 on a scale of 1 (very unsure) to 7 (very sure).

Given the excellent performance of adults, the performance of the preschool group was of particular interest. The results were strikingly similar to those of adults. Four-year-olds based 68% of their inferences on category membership despite the lack of perceptual support for such inferences. This figure was significantly above chance and significantly higher than the results of a control condition in which subjects were shown the labeled test pictures only (e.g., shark) and were asked the test questions (e.g., whether it breathed underwater or above the water). Moreover, 37% of the children consistently based their inferences on category membership across items, whereas none of the subjects consistently based their answers on appearances.

These results were replicated with younger children in a follow-up study. Gelman and Markman (1987) examined patterns of inferences in 3- and 4-year-olds using a slightly different procedure. For each of 10 picture sets, the task involved teaching subjects about a novel property of a single target picture (e.g., “See this cat? It can see in the dark”) and then pairing the target with four test pictures in turn. The children were then asked whether the property was true of each test picture. There were four types of test pictures: (a) same category, similar appearance (e.g., a cat with markings similar to those of the target cat), (b) same category, dissimilar appearance (e.g., a cat with different markings, in a different position), (c) different category, similar appearance (e.g., a skunk with markings strikingly similar to those of the target cat), and (d) different category, dissimilar appearance (e.g., a dinosaur).

The experiment had three conditions designed to tease apart the effects of language and appearance. In the word and picture condition, all pictures were shown and named. In the word-only condition, all pictures were named but the target picture was not shown to the child. Instead, the experimenter simply said, for example, “I see a picture of a cat. This cat can see in the dark . . .” In the picture-only condition, all pictures were shown, but none were named.

The results support the view that children rely on natural kind category membership to guide their inferences (Figure 5.1). In the word and picture condition, the children drew more inferences to members of the same category with dissimilar appearances (64% overall) than to members of a different category with similar appearances (29% overall). For example, when told that a small brown snake lays eggs, they were more likely to infer that a large gray cobra lays eggs than to infer that a small brown worm lays eggs. In contrast, there were important within the target category: The children drew more inferences to similar members of the target category (e.g., another small brown snake) than to dissimilar members (e.g., a large gray cobra). However, no such trend was apparent outside the target category (the children drew as many inferences to a cow as they did to a small, brown worm).

As expected, in the word-only condition, the children also based their inferences on category membership and not appearances. It was only in the picture-only condition that they were more apt to draw inferences based on appearances. They drew more inferences to pictures similar in appearance to the target but from a different category (e.g., from a tan shell to tan stone of the same shape) than in the word and picture condition. And for pictures that differed in category membership from the target, they drew more inferences to those pictures similar in appearance (e.g., from a black and white striped cat to a skunk) than to those dissimilar in appearance (e.g., from a black and white striped cat to a dinosaur). Few significant differences were found between 3- and 4-year-olds, and none were found in the word and picture condition. By age 3, children use information about category membership to guide inferences.

In fact, children as young as age 2 are sensitive to the rich infer-
The acquisition of natural kind terms

Figure 5.2. Sample item set from Gelman and Coley (1990).

The potential power of natural kind categories (Gelman & Coley, 1990). Children ranging in age from 2.1 to 3.1 (mean age 2.8) were tested on a simplified version of the inference task already described. Specifically, for each of nine items, children were shown a target picture from a familiar category (e.g., a bluebird) and reminded of a familiar property that described it (e.g., "This bird lives in a nest"). This target picture was then set aside and not referred to again, but was left in view. Each subject was then shown four test pictures, one at a time. These included (a) a typical instance of the target category (another bluebird), (b) an atypical instance of the target category (a dodo), (c) a typical instance of a contrasting category (a stegosaurus), and (d) an atypical instance of the same contrasting category (a pterodactyl) (see Figure 5.2 for a sample set). Furthermore, the atypical instance of the contrasting category was perceptually quite similar to the target picture, whereas the atypical instance of the target category was not (e.g., the pterodactyl looked like the target bluebird, but the dodo did not). Perceptual similarity was confirmed by adult ratings. Finally, for half the subjects, category membership was conveyed via verbal labels (e.g., "This is a bird"); the other half heard no labels.

As expected, subjects performed well on the typical items in both conditions (overall, 75% correct). For example, the subjects reported that the bluebird lived in a nest and the stegosaurus did not, whether or not these pictures were labeled. For the atypical items, however, the subjects who heard the labels drew correct inferences at a level significantly above chance (69% correct), whereas those who did not hear a label performed significantly below chance (42% correct). For example, the children inferred that a dodo, labeled bird, lived in a
nest like the target bird, but that a pterodactyl, labeled *dinosaur*, did not. When pictures were not labeled, 2-year-olds were more likely to infer that the pterodactyl lived in a nest and that the dodo did not. These results also held up when the performance of individuals was examined. In the label condition, 6 of 11 children consistently based their inferences on category membership (i.e., they answered on the basis of category membership for 26 or more of the 36 trials); none based their inferences consistently on appearances. In the no-label condition, one child responded on the basis of category membership, whereas 6 of 11 consistently answered on the basis of appearances.

Thus, for children as young as age 2, category membership promotes important inferences. This is a remarkable achievement, given the age of the subjects, their lack of formal scientific training, and their otherwise persistent attention to salient perceptual features.

*Control studies.* Because category membership in these studies was conveyed by language, children's performance could conceivably have been due to a simple "same word—same property" strategy. That is, the children may have had a response bias to say that category members were alike just because the experimenter had labeled them with identical words. This seemed unlikely, because the task required the children to say the property explicitly (not just to point), and because the children gave many sensible justifications (Table 5.1). Nonetheless, several control studies, described later, show that the children were not answering solely on the basis of hearing the same word for two pictures.

In some cases, children drew appropriate categorical inferences without identical labels. Four-year-olds performed well when category information was conveyed by synonymous rather than identical labels (e.g., the target was referred to as a "puppy" and the test picture was referred to as a "baby dog"; Gelman & Markman, 1986). Children in this condition based their inferences on category membership 63% of the time, which is significantly above chance.

Children sometimes drew category-based inferences even in the absence of labels. In the no-label condition of Gelman and Markman (1987), in which none of the pictures was named, children occasionally figured out category membership on the basis of subtle perceptual clues and then proceeded to draw correct inferences (3-year-olds on 3 out of 10 items and 4-year-olds on 5 out of 10 items). For example, the children did not draw inferences from a leaf insect to a leaf—despite a striking similarity between the two—because they attended to subtle features (e.g., the leaf insect had eyes and antennae, which the leaf lacked). When a separate group of children was later asked to name all the pictures used in the study, correct naming correlated highly (r = .77) with the children's ability to draw category-based inferences. Those pictures that the children named correctly (e.g., the leaf insect, which was typically called "a bug") were the ones for which the children also drew appropriate inferences.

In other cases, identical labels did not yield unwarranted inferences. In one control condition (Gelman & Markman, 1986), children were presented with a meaningless task for which category membership was irrelevant. In this task, the experimenter labeled and placed a different-colored chip on each of the two target pictures in each set. (The picture sets and labels were identical to those used in the induction task described earlier.) The children were then asked which color should go on the test picture. For example, if the tropical fish received a red chip and the dolphin received a blue chip, should the shark (labeled *fish*) receive a red or a blue chip? If children have a response bias to match pictures on the basis of the labels, the results should be identical to performance in the induction task. However, if performance in the induction task reflected the belief that natural kind categories share deep similarities, chil-

<table>
<thead>
<tr>
<th>Table 5.1. Sample justifications of category-based choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>A squirrel eats bugs, &quot;because it's a squirrel.&quot;</td>
</tr>
<tr>
<td>A fish stays underwater to breathe, &quot;because he's a fish.&quot;</td>
</tr>
<tr>
<td>The gold nugget melts like the gold bar, &quot;because they're both the same thing.&quot;</td>
</tr>
<tr>
<td>A chunk of salt helps make snow melt, &quot;because it's the same kind as this [fine-grained salt].&quot;</td>
</tr>
<tr>
<td>A bug breathes air in, &quot;because some bugs work like this.&quot;</td>
</tr>
<tr>
<td>A flower has tubes for water inside, &quot;because every flower has tubes inside, so it does have tubes inside!&quot;</td>
</tr>
<tr>
<td>The coral has to catch food, &quot;because all corals do that.&quot;</td>
</tr>
<tr>
<td>The dinosaur has cold blood, &quot;cause every dinosaur has cold blood, even when it's frozen.&quot;</td>
</tr>
</tbody>
</table>
dren would have no basis for deciding which color chip to select and performance on the control task should be at chance, with subjects choosing each color about half the time. Indeed, children’s performance in this condition was not significantly different from chance.

Similarly, in another control condition, when 4-year-olds learned properties that should be more generalizable on the basis of appearance than category membership (e.g., “weighs 10 pounds”), they did not base inferences on categories over appearances (Gelman & Markman, 1986). Likewise, 2-year-olds who heard pictures labeled with adjectives describing temporary states (e.g., sleepy) instead of category names (e.g., dinosaur) performed just like those who had heard no labels at all (Gelman & Coley, 1990). See S. A. Gelman (1988) for further evidence that children are more likely to generalize enduring biological properties concerning, for example, internal structure, diet, or function (e.g., “has parts made out of calcium,” “likes to eat alfalfa,” “needs branches to breathe”) than transient properties dealing with temporary states or historical accident (e.g., “smells yucky,” “fell on the floor this morning,” “is a year old”).

Thus, children do not blindly base their inferences on the similarity of words that are associated with each picture. Not all words promote categorical inferences, and not all categorical inferences require labels. Rather, children use category labels to guide inferences about important underlying properties.

Role of theories. Carey (1985) argues that a naïve theory of biology best explains children’s patterns of inferences across basic-level animal categories. Rather than look at patterns of induction within basic-level natural kind categories, Carey has investigated how children project properties taught about one basic-level animal category to other categories differing in similarity to the target category. Subjects were taught that people, dogs, or bees had a spleen (“a green round thing”) inside of them. When taught that people had spleens, young children inferred that other animals had spleens on the basis of their similarity to people. Mammals were most likely to have spleens, down through worms, which were the least likely. However, when children were taught that dogs or bees had spleens, the property was not generalized on the basis of similarity to dogs or bees. For instance, children generalized from people to aardvarks 76% of the time, but from dogs to aardvarks only 29% of the time. Similarly, children generalized from people to stinkbugs 52% of the time, but from bees to stinkbugs only 12% of the time. These results run counter to what one would expect if children were basing their inferences on similarity alone, since, for example, bees and stinkbugs are more similar to one another than are people and stinkbugs. Nonetheless, the findings can be explained if 4-year-olds hold the belief that people are the prototypical animals and thus the prototypical possessors of animal properties. If 4-year-olds hold this theory, then animal properties taught about people should be readily generalizable, whereas those taught about more peripheral animals (dogs and bugs) should be less so. Further support for this interpretation is that children generalized from dogs to people only 18% of the time, whereas they generalized from people to dogs 71% of the time. Carey argues that “the prototypicality of people plays a much larger role in determining 4-year-olds’ projection of having a spleen than does similarity among animals” (p. 128).

In a related experiment, Carey (1985) found that subjects as young as age 7 rated a mechanical toy monkey more similar to a person than to any of a number of pictured animals. However, these subjects rarely (12% of the time) attributed animal properties such as eating, sleeping, and having babies to the toy monkey. In fact, they attributed animal properties significantly less often to the mechanical monkey than they did to a worm, which was understandably rated as very dissimilar to a person. By age 7, children are aware than people and worms are alike in important ways that people and mechanical monkeys are not, despite similar appearances. Overall, Carey’s work demonstrates that children do not base inductions across categories on perceptual similarity alone. Rather, such inductions are informed by children’s emerging biological theories.

Taken together, these results demonstrate that children are indeed sensitive to the importance of natural kind categories for promoting inferences beyond what can be perceived. By 2 years of age, language signals category membership; and categories promote rich inductions beyond what is observable. This is true even in the strongest case, in which category membership receives little perceptual support.

Knowledge of nonobvious properties

Nonobvious properties are especially important for natural kind categories. Evidence from several domains points to the conclusion that
children are knowledgeable about nonobvious properties and realize their importance for the way objects function.

Understanding of insides. Children's knowledge of unobservables is demonstrated by their relatively accurate descriptions of the insides of objects. Gelman and O'Reilly (1988) asked preschool and second-grade children whether a range of categories (animals, plants, vehicles, clothing, etc.) had the same insides. For example, children were asked whether each of the following had "the same kinds of stuff inside": (a) all dogs, (b) dogs and horses, (c) dogs and snakes, and (d) dogs and tractors. They were also asked to describe what each had inside. In explaining their answers, the children appropriately mentioned internal parts 65% of the time and external parts or functions only 12% of the time. Furthermore, on natural kind items children discussed natural kind parts (liquids, organs, bones) 35% of the time, whereas they mentioned artifact parts (stuffing, metal, fabric) only 1% of the time. This pattern was reversed for artifact items: The children appropriately mentioned artifact parts 59% of the time and mentioned natural kind parts only 4% of the time. Although not always accurate in detail ("Carrots have milk inside. Milk and wheat; "Teddy bears have feathers inside"), the subjects knew what kinds of insides were appropriate for different kinds of items.

Similarly, R. Gelman (1987) asked preschool children what was on the inside and on the outside of various animate and inanimate objects, including humans, elephants, cats, dolls, and puppets. As in Gelman and O'Reilly (1988), children distinguished animals from inanimate objects on the basis of internal properties. They tended to say that animals had blood, bones, or other internal organs on the inside, whereas inanimate objects had material, mechanical devices, or nothing inside. Moreover, children seemed to believe a more general rule, that the inside of an animal causes its self-generated movement ("causal innards principle"). As evidence for this principle, R. Gelman found that children expected the insides of different animals to be very similar to each other and to differ from their outsides. In contrast, for inanimate objects children often reported that the insides were the same as the outsides (a "surface generalization rule"). For example, children might say that a doll has material on both the outside and the inside, whereas an elephant has skin outside and blood inside. Again, preschool children differentiate between animals and inanimate objects on the basis of nonobvious properties.

Children can also attend to internal composition when it conflicts with surface appearances. Gelman and Wellman (in press; cited in Gelman & Coley, 1989) showed children triplets of items in which the target item and one test item looked alike but were from different categories (e.g., a pig and a piggy bank), and the target item and the other test item did not look alike but were from the same category (e.g., the pig and a cow). For each pair of test items, children were asked two questions: in this example, "Which of these looks most like the pig?" and "Which of these has the same kinds of insides as the pig?" These questions required the subjects to switch their answers; the first required attention to surface similarity, and the second to the less obvious property of internal makeup. Four-year-olds performed better than 3-year-olds (78 vs. 58% correct), but both performed better than chance. Gelman and Wellman also conducted an analysis of error patterns that took into account how the children answered both questions for a given triplet (i.e., both the "looks like" and "has insides" questions). It is interesting that on this analysis children erred as often by saying that things that had the same insides not only had the same insides but also looked alike (e.g., the pig and the cow looked alike) as by saying that things that looked alike not only looked alike but also had the same insides (e.g., the pig and the piggy bank had the same insides). These results show that the children were able to consider both appearances and nonobvious properties in the same task and that not all errors were due to emphasizing outward similarity over category membership.

Children are also able to sort objects with identical outward appearances on the basis of internal, hidden (but remembered) characteristics. Deloache and Todd (1988) found that 5-year-olds could sort identical closed containers into two groups of six, depending on whether they had seen M&M candy or pegs put into each container before it was closed. Overall, the children succeeded with this sorting task on 78% of the trials. In a follow-up experiment, 5-year-olds were asked to form similar categories, but in this case two different-looking types of containers were used. Half of each type of container contained candy and half contained pegs. To succeed in this task, the subjects not only had to attend to nonobvious properties (i.e., remember which containers held which objects) but also
had to ignore the salient appearances of the containers. Although this task was more difficult than the previous one (44% successful sorting overall), by the fourth trial the subjects were able to sort successfully on the basis of contents 62% of the time.

Even preverbal infants can attend to nonobvious properties. In another categorization task, Kolstad and Baillargeon (1989) examined infants’ ability to categorize on the basis of both perceptual and functional properties. To do this, they habituated 10½-month-old infants to series of events in which three brightly colored cylindrical containers were manipulated in a viewing box. In the experimental condition, subjects watched salt being poured into and out of the containers; in the control condition the movements were the same, but no salt was used. After habituation, the infants’ looking times were recorded for two test events, either with a container that resembled the training containers but apparently had no bottom (a “tube” that actually had a transparent bottom) or with a floral-pattern box that did not resemble the training containers but did have an obvious bottom. Each test container was manipulated either with salt (experimental condition) or without salt (control condition). In the experimental condition, both containers actually held the salt. Infants in the experimental (salt) condition looked longer at the event involving the “bottomless” tube that held salt than at the event involving the box that held salt. Infants in the control (no salt) condition looked longer at the box event. The authors argue that in the experimental condition, the subjects focused on containment and therefore categorized the box with the other containers. The infants were surprised that the apparently bottomless tube actually held the salt and so gazed longer at that event. The control children had no reason to be surprised and therefore focused on perceptual similarity, classifying the tube with the other brightly colored cylinders and seeing the box as novel. Children as young as 10½ months are not limited to categorization on the basis of perceptual similarity alone.

Understanding the animate–inanimate distinction. Massey and Gelman (1988) demonstrated that 3- and 4-year-old children can make a conceptual distinction between animals and inanimate objects on the basis of subtle perceptual cues. Children were shown photographs of unfamiliar objects from five categories—mammals, nonmammalian animals, lifelike statues of mammals, wheeled vehicles, and complex rigid objects—and were asked whether each object “could go down (up) a hill all by itself.” The statues and the mammals looked more alike than did the mammals and the nonmammalian animals (e.g., praying mantis, tarantula). Overall, 4-year-olds were correct on 90% of the trials, and 3-year-olds on 78%. Twelve of 20 subjects reliably said that animals, but not inanimate objects, could go up and down a hill by themselves. The correct use of this “animacy” rule required attention to subtle cues rather than overall similarity. Five more subjects used a variant of the animacy rule, but sometimes denied that an animal could go up and down a hill by itself for practical reasons (e.g., the hill was too big). The authors concluded:

Performance on this task cannot be attributed to simple perceptual prototypes or rules that are based on single perceptual features. Clearly, the children were using perceptual information, because pictures were their only direct source of information about these unfamiliar objects. However, the patterns in their answers are not predicted by the obvious perceptual similarities and dissimilarities among the items. (p. 316)

This finding is consistent with other work showing that children are knowledgeable about the distinction between animate and inanimate objects from an early age (e.g., Dolgin & Behrend, 1984; Gelman, Spelke, & Meck, 1983; Richards & Siegler, 1984, 1986).

Understanding of mind. Children’s understanding of the mind is another realm in which an appreciation for unobservables is apparent. Wellman and Gelman (1988) note at least two ways that children’s theory of mind demonstrates an appreciation for the nonobvious. First children, like adults, distinguish between internal thought and external behavior. Wellman and Estes (1986) showed that children as young as age 3 correctly report that physical objects—and not thoughts or dreams of such objects—can be touched and manipulated, or seen by someone else. Second, children are aware that internal psychological states offer the best explanation of others’ external behavior. For example, Wimmer and Perner (1983) show that 4-year-olds can correctly predict that a boy in a story will look for a cookie in a drawer because he believes the cookie is there, despite the fact that the subjects know the cookie has been moved to a cupboard. Through their differentiation between thoughts and objects and their use of beliefs to predict behavior, preschoolers show an understanding of the nonobvious.
In this section, we have shown that children's understanding in a variety of domains, including knowledge of objects’ insides, animacy, and the mind reveals that children are aware of, and are able to attend to and use, nonobvious properties.

Psychological essentialism

Beyond appreciating unobservable but important properties that category members have in common, children also believe that members of a category share some unique underlying property that is responsible for the other similarities associated with category membership.

One way that this psychological essentialism can manifest itself is through a belief in innate potential. Recognition that an animal has the potential to develop into an outwardly different mature form (e.g., recognition that a tadpole has the potential of becoming a frog) is consistent with belief in a category essence. Recent work suggests that 42% to 5-year-olds believe that individuals have an innate potential, even before that potential visibly manifests itself (Gelman & Wellman, in press; cited in Gelman & Coley, 1989). In this study, children were shown an immature member of an animal or plant category (either a baby animal or a seed) and then were told that the baby or seed grew up in an environment better suited to another species (e.g., a cow that grew up on a pig farm and never saw another cow, or a seed that came from an apple but was planted in a flower pot). Next, the subjects were asked questions about what the baby animal or seed would be like when it grew up. Overall, 80% of the time children answered that the baby would have the characteristics expected given its category membership, despite the conflicting evidence. For example, children answered that the baby cow would moo and have a straight tail when it grew up and that the apple seed would grow into an apple tree. These data provide preliminary evidence that children are essentialists. They assume that members of a category share an early-emerging (or innate) potential that can overcome a strong environmental influence.

Another way to examine psychological essentialism in children is to ask them to explain why objects have certain properties. If children explain such phenomena by appealing to intrinsic factors rather than external, imposed factors, they are showing indirect evidence of essentialism. Kremer and Gelman (1989) asked children questions such as “Why do rabbits have long ears?” and “Why do birds fly?” They found that 72% of 4-year-olds and 73% of first-graders spontaneously mentioned inborn dispositions, intrinsic nature, or growth at least once. For example, preschoolers answered that a rabbit has long ears because “the egg made the [rabbit’s] ears so that it had them when it hatched,” or birds fly “because that’s the way birds are made.” Moreover, children mentioned these factors significantly more often when explaining properties of natural kinds (e.g., rabbits, flowers, salt) than when explaining properties of human artifacts (e.g., cars, crayons, phones). Apparently, children appeal to essential properties such as inborn factors, intrinsic nature, and growth to explain why natural kind category members share certain attributes.

Essentialism also entails the belief that deeper properties are more important for category membership than are superficial ones. Nelson (1974) provides arguments and evidence that in the earliest stages of word learning children look for defining qualities of objects that extend beyond obvious similarity features. More recently, Keil (1986) examined these issues in older children. He presented children in kindergarten, second grade, and fourth grade with stories about animals that had appearances and habits characteristic of one natural kind (e.g., raccoons), but upon closer scrutiny (e.g., inspection through microscopes) were found to have a cluster of features of another natural kind (the raccoonlike animals actually had the internal organs of skunks, their parents and babies were skunks, etc.). Five-year-olds tended to maintain that the creatures were raccoons. Second- and fourth-graders and adults were willing to let discoveries override appearances; they tended to say that the creatures were skunks.

The relatively poor performance of 5-year-olds on this task seems to run counter to the claim that preschool children are psychological essentialists. However, Keil's task involved inferring categories from properties (e.g., “This animal has properties x, y, and z. Is it a skunk?”). Young children have been shown to find this sort of inference task much more difficult than the converse task of inferring properties from categories (e.g., “This animal is a skunk. Does it have properties x, y, and z?”), at least with gender categories (Gelman, Colman, & Maccoby, 1986). The performance of the younger subjects could reflect this difficulty rather than a lack of essentialism.
Another way to examine childhood essentialism is to ask children about transformations of objects. If an object has an essence that determines its category membership, then superficial changes in that object that do not affect its essence should not affect its category membership. Keil (1986, 1987) reports that by age 5, children correctly report the identity of an object when superficial transformations cross ontological boundaries. For example, when asked about a porcupine that became cactuslike in appearance (dyed yellowish green and injected with a substance that caused it to hibernate for years), 5-year-olds answered that the animal in question was still a porcupine. Results were different for transformations occurring within ontological boundaries. Five-year-olds tended to say, for example, that a raccoon that was shaved, painted black and white, and implanted with an odor sac was now a skunk. Reliance on appearances decreased with age; children became more essentialistic as they grew older. By second grade, children preserved category membership within ontological boundaries on a majority of their judgments. Again, the errors of younger children could be attributed to their difficulty in inferring the category of an object given properties of that object rather than a lack of essentialism.

Recent work hints at the kinds of properties children consider essential. Springer and Keil (1989) found that children predicted that unrelated members of an animal category would share abnormalities that were inborn, but not abnormalities that were acquired. For example, children reasoned that a bull was more likely to be born with a pink heart inside if other bulls that lived nearby had been born with pink hearts inside than if other bulls had a one-time accident that made their hearts pink. Children are aware that inborn features are more likely to be intrinsic to species, and therefore essential.

Although more work must be done, these studies offer some evidence that children approach concept learning with at least the roots of psychological essentialism.

Acceptance of anomalies

By 2 years of age, children have a rudimentary ability to accept anomalies that are explicitly labeled. In the studies reported earlier, nearly all the children accepted the anomalies presented (e.g., a long-eared squirrel or a birdlike flying fish). They occasionally commented on

the anomaly (e.g., “That’s a funny looking squirrel”), but generally they accepted it (Gelman & Coley, 1990; Gelman & Markman, 1987).

Nonetheless, children may have difficulty accepting category anomalies that are not explicitly pointed out or explained to them. We predict that anomalies would be difficult for children because they conflict with the belief that category members have much (including appearances) in common. In one study, 4-year-olds accepted anomalies that adults considered possible, such as “a rabbit that doesn’t like carrots,” only 26% of the time (Biderman, 1969).

By age 6, children accepted such anomalies 53% of the time. (Adults accepted these “possible anomalies” 91% of the time.) Even when trick photographs depicting these anomalies were shown to the children, preschoolers still accepted fewer anomalies than 6-year-olds (55 and 90%, respectively). At both ages, children consistently distinguished between possible anomalies (such as the rabbit just mentioned) and impossible anomalies (such as “a whale that lives out of water”). Nonetheless, they were rather conservative in their judgments of whether anomalous category members were possible.

Overall, it appears that children have some ability to accept anomalous category members when presented with them, but otherwise tend to discount the possibility of their existence. Children learn to accept anomalous category members later than they learn that categories promote inferences, share nonobvious properties, and have essences.

Division of linguistic labor

There is informal evidence that children may recognize a division of linguistic labor in the meaning and use of category terms. In our previous work (e.g., Gelman & Markman, 1986, 1987; Gelman & Coley, 1990) children accepted the labels provided by the adult experimenter even when they were inconsistent with what they initially thought the names to be. One 2-year-old subject even articulated this principle to his mother after participating in one of our experiments. When discussing a sticklike snake shown in one picture, he said, “I thought it was a stick, but the man [i.e., the experimenter] said it was a snake.” This initial understanding that linguistic knowledge may be greater among a subset of the community could arise out of children’s early grasp of the asymmetry between children and adults, in both knowledge and authority. Unfortu-
nately, to date we still lack any rigorous investigation of the development of this understanding.

Corrigibility

Theory-laden categories are subject to revision on the basis of new information. If children appreciate the corrigibility of categories, they should be willing to revise initial categorizations on the basis of new discoveries. This ability would include the capacity to accept category anomalies that were originally misclassified. We know of little direct evidence, but we expect this capacity to be late-developing. Corrigibility requires appreciating the limits of one’s categories and contrasts with properties such as essentialism that stress the importance of categories for revealing hidden realities about the world. To the extent that one believes a category to be richly structured, nonarbitrary, and reflecting the natural order, it should be difficult to abandon, question, or restructure that category. Moreover, corrigibility implies flexibility, and there is some evidence that early cognitive structures may be relatively inflexible (Inhelder & Piaget, 1964; Keil, 1986, 1987). To examine the issue further, we need studies that provide children with new information that is inconsistent with their present category knowledge and to examine how readily children’s categories change to incorporate the new information.

Summary

In this section we have presented evidence that children’s natural kind concepts exhibit many of the properties we outlined for adults’ natural kind concepts. Specifically, they promote rich inductive inferences, embody knowledge of the importance of unobservable properties, reveal the roots of psychological essentialism, and make some allowance for anomalous members. In the next section, we will outline our proposal for how children’s categories come to possess these qualities.

The role of language in children’s construction of natural kinds

We have shown that children’s natural kind categories exhibit many of the same properties as the natural kinds of adults, although children have little in the way of explicit scientific training. The implication is that children do not require expertise, scientific training, or even an explicit theory to assume that categories are richly structured beyond surface similarity. How is this accomplished?

For adults, there are at least two ways that concepts may go beyond appearances without explicit theory. First, extensive experience can direct one’s attention to nonobvious yet informative perceptual cues or feature correlations. For example, expert chick sexers learn to distinguish between male and female day-old chicks on the basis of years of experience classifying and getting feedback from experts (Biederman & Shiffrar, 1987). They do not have an explicit biological theory to guide their classification. Rather, they learn the exceedingly subtle perceptual cues by extensive exposure and feedback. Given children’s lack of extensive experience, perceptual learning does not offer a promising account of children’s early appreciation of natural kinds, although it probably plays an increasing role in later development.

Second, conceptual change can be induced in adults without theory via language, without extensive knowledge or perceptual learning. For example, suppose that you initially believe that bamboo plants are members of the category tree. On learning that bamboo is actually a grass, you may modify your conceptual structure to classify bamboo as having deeper properties in common with other grasses than with trees (e.g., stem structure and rate of growth). This change would not necessarily be accompanied by any theoretical understanding of why bamboo is a grass. Nor would it necessarily result from learning subtle perceptual distinctions – via either theory or pure experience – that betray bamboo’s true identity. Rather, the label grass conveys the category identity. The assumption is that certain named categories capture deep similarities despite outward appearances and despite the nature of one’s particular theory.

To what extent do children use language to help construct natural kind categories? There are two extreme models that may characterize the role of language in children’s understanding of natural kinds. According to a conceptual position, language has no role in the way children organize their knowledge and experience; at most it reflects a conceptual understanding, but it does not shape it. In contrast, a linguistic position holds that language is the primary source of information about the structure of the world. On this view, children assume that things called by the same name are natural kinds, shar-
ing deep similarities and a category essence. Words are primary; concepts are molded to fit the language.

Neither extreme position can account for the available data. We have already seen that providing a label can alter the way children reason about an object, thus arguing against the conceptual position. For example, calling a pterodactyl a “dinosaur” leads children to infer that it does not live in a nest, whereas without the name children assume that it does (Gelman & Coley, 1990). At the same time, words sometimes fail to promote inferences for children (see, e.g., the adjectives used in Gelman & Coley, 1990), thus arguing against a strong linguistic position. Indeed, there are many words in the adult language that children must at some point learn to distinguish from natural kind terms, including homonyms (but), superordinates (furniture), simple adjectives (striped), and metaphors (mouse).

We hypothesize that children have several expectations that jointly determine when a word maps onto such a richly structured category. Without evidence to the contrary, children will assume that an object word maps onto a natural kind category. They will make this assumption as long as the set of objects named by the word is reasonably coherent (excluding, e.g., red things). Once children determine that a word names a natural kind, this conclusion is fairly robust and will persist even when they encounter anomalies that would seem to contradict the grouping. The coherence of the category plus the belief in hidden underlying similarities allows children to accept anomalies that are explicitly presented. For example, once a child decides that bird maps onto a natural kind category, when learning that a dodo is a bird the child will assume that it is the same kind of thing as other birds (an otherwise coherent category consisting of robins, sparrows, etc.). Thus, language is critical for reorganizing or restructuring categories away from groupings based on similarity alone. Objects that might otherwise not be considered members of a category will be classified (or reclassified) once the name is learned. In contrast, if a category lacks coherence when it is initially learned, then a child will not assume that the new word maps onto a natural kind. For example, a child will not consider mammal a natural kind term if it labels a diverse group including aardvarks, wallabies, bison, and ferrets.

Thus, children’s expectations can be viewed as an interplay of three distinct assumptions. On this account, children’s understanding of natural kinds is both linguistically and conceptually based. We are not saying that language determines the categories children have. Rather, children have certain expectations about categories and how they are named. These expectations together influence the nature of the categories a child constructs.

**Hypothesis 1: Natural kind assumption**

First, we propose that children assume that every object is a member of some natural kind category (i.e., with an underlying essence, potential to promote inferences, nonobvious properties, etc.). This can be thought of as a theory placeholder view analogical to the essence placeholder notion of Medin & Ortony, 1989). The claim is that children have a broad expectation that objects will fall into kinds and that these kinds will share theoretically important properties whether or not those properties are currently known. The expectation that categories will have a deeper theoretical basis may precede the actually filling out of the substance of the theory. Hearing a name for the first time should be enough evidence for a child to infer that it is a natural kind term. However, this assumption is also possible without language and may operate prelinguistically (see Baldwin, Markman, & Melartin, 1989, for evidence that infants make inductions from unnamed categories).

**Evidence.** As detailed in the preceding section, preschool children treat a wide range of categories as having an essence, rich inductive potential, and nonobvious features. Even wholly novel categories can serve as the basis for induction (Davidson & Gelman, 1990). Evidence that children hold a general assumption that objects fall into natural kinds comes from the finding that children overgeneralize the inductive potential of categories. Recall that some categories — particularly simple artifact categories (e.g., chairs, hammers) — have few if any theory-laden properties. We would expect artifact categories to promote many fewer inferences than natural kind categories (see S. A. Gelman, 1988, for a discussion). Indeed, elementary school children and adults draw many fewer inferences within artifact categories (e.g., televisions or cups) than within natural kinds (e.g., rabbits or flowers). However, preschool children tend not to distinguish them, at times overgeneralizing their inductions from artifacts (S. A. Gelman, 1988; Gelman & O’Reilly, 1988). For ex-
ample, preschool children are as likely to infer that different chairs have the same internal substance as they are to infer that different dogs have the same internal substance. It seems that young children have a general assumption that all objects belong to natural kind categories, an assumption that gets refined with age.

Hypothesis 2: Linguistic transparency

We hypothesize that children assume that every natural kind has a name and that names convey category membership and all the properties that go along with category membership. Coupled with the natural kind assumption, this implies that certain words may function like Kuhnian paradigms: "object[s] for further articulation and specification under new or more stringent conditions" (Kuhn, 1970, p. 23). That is, a word can serve to stake out a new category, which then must be explored in more depth, a "lure for cognition" (Brown, 1956).

Although there is no direct evidence that children assume that all natural kinds are encoded in language, clearly they do assume that at least some natural kinds are expressed in words. This is evident in the way words modify children's conceptions of the objects being labeled. As shown earlier, children draw inferences on the basis of familiar category labels. Without a label, they may decide that a dodo does not live in a nest; if they are told that it is a bird, they declare that it does. In a very real sense, then, our natural kinds are passed down to children through the language we speak. The naming practices of the culture are fundamental to the content of our natural kinds.

Hypothesis 3: Coherence

Children appear to assume that membership in a natural kind category is typically reflected in outward signs (i.e., those properties that are immediately apparent, including appearances and salient behaviors or traits). Objects that are more alike in a Roschian sense (Rosch, 1978) are more likely to share an essence. So, for example, objects with similar shapes, parts, functions, and textures would, on the child's first guess, probably belong to the same natural kind category. The implication is that basic-level categories are likely to be natural kinds. At the same time, however, similarity is at best only a fallible guide to natural kind identity. Thus, similarity serves as a heuristic rather than a defining basis.

It is important to note that coherence operates on a category taken as a whole, not on individual instances. Thus, if trying to determine whether to draw an inference from Object A to Object B, a child would not simply calculate the similarity between the two objects. Rather, the child would determine whether A and B belong to members of the same natural kind category, using as a guide the coherence of the category that encompasses both A and B. If A and B receive the same name, the child could start by calculating the coherence of the category that receives that name, to determine whether it is a natural kind. Thus, even a very anomalous object could be a member of a highly coherent category (e.g., dogs may be considered highly coherent, even if a child also comes across a chihuahua). Or if two objects are highly similar but are known to belong to different natural kinds (e.g., a bird and a pterodactyl), children will not try to calculate the coherence of the grouping that includes both birds and pterodactyls and so will not assume that they have underlying properties in common.

Evidence. Children are aware that more coherent categories provide a better basis for induction than less coherent categories. Within a hierarchy, basic-level categories (e.g., chairs) are more coherent than superordinate-level categories (e.g., furniture) (see Waxman, Chapter 4, this volume). In one study (Gelman & O'Reilly, 1988), preschoolers, second-graders, and adults drew many inferences at the basic level and significantly fewer inferences at the superordinate level. When taught a new property of a familiar object (e.g., "This dog has leukocytes all through it"), they readily drew inductions to members of the same basic-level category (e.g., inferring that other dogs also have leukocytes inside). However, they drew few inferences to members of the same superordinate-level category (e.g., inferring that other animals - horses and snakes - have leukocytes inside) even when they were all explicitly labeled animals. (All items were pretested to make certain that even the youngest subjects knew both the basic- and superordinate-level names of these pictures.) For example, children tended to deny that a horse or a snake would have leukocytes inside like the dog, even when all three were repeatedly called animals.

There were also developmental differences, with preschoolers
showing the most dramatic difference between basic and superordinate level. Thus, preschoolers showed the strongest tendency to distinguish coherent from less coherent categories. These results are consistent with the results of Carey (1985) and Inagaki and Sugiyama (1988), who also found few superordinate-level inferences (from the categories animals and living things) but did not provide labels for these categories. It may be that superordinate categories are more coherent for older children and adults than for younger children (Murphy & Medin, 1985).

More evidence for the importance of coherence comes from a series of studies by Davidson and Gelman (1990) examining children’s inductive inferences within novel categories such as gnulike animals with trunks. The major finding was that category labels affected whether children drew inferences, but only when the words mapped onto coherent categories. There were two ways this could happen. First, a category could inherit coherence by means of language, when a set of novel pictures was named by a familiar natural kind term. For example, children drew many more inferences from one object to another when both were called “cows” than when both were called “zavs.” Presumably, the category cow is already coherent, whereas the category zav is not. Second, a category gained coherence when there was some correspondence between the way sets of objects were named and their appearances. When naming and perceptual similarity were completely orthogonal to one another (e.g., when “zavs” were as dissimilar from each other as they were from “feps”), children ignored the labels and based their inferences strictly on appearances. However, when there was even a slight correlation between naming and appearance (“zavs” were more similar overall than were “zavs” and “feps”), children used both the labels and appearances as the basis of their inferences. It was not the similarity of pictures taken individually that determined whether children used a word as the basis of induction, but rather the coherence of the set of pictures given the same name. Thus, conceptual grasp of the category being named, and not just the label itself, influences children’s expectations about a label.

In short, we propose that children start with an assumption about categories, that they are natural kinds and that natural kind categories are typically signaled by appearances. Children then look for words that express these categories. Once a word is found that expresses a natural kind, it can help organize new information, promote inferences, and so on. Note that children do not have a completely general assumption about language per se, but rather about how language and categorization interact.

**When is a category coherent?**

There is a critical interplay between coherence and linguistic transparency. Consider category anomalies, such as penguins. We might not at first include penguins within the bird category, but given the appropriate linguistic information we are willing to stretch the category to include them. So language can identify members of a category when appearances would not. Yet if the amount of stretching required is too great (e.g., stretching the bat category to include both flying bats and baseball bats), then linguistic transparency is overcome by lack of coherence. A crucial question that remains concerns how the difficult cases get decided — how powerful is language, how powerful is conceptual coherence? At the least, we know that language can stretch (and compress) a category beyond what its boundaries would be on a purely nonlinguistic basis.

At this point we are just beginning to understand what category coherence consists of. However, it seems clear that three factors in particular could be especially useful in helping children determine whether a category is coherent: language form, parental input, and prototypes.

**Language form.** Language form (“part of speech”) can give children a clue as to whether the category named is theory-laden. In particular, common nouns are much more likely to refer to theory-laden categories than are adjectives (see Markman, 1989, for a more detailed discussion of this possibility). If children are sensitive to linguistic form, then simply by knowing that white is an adjective and not a noun, a child could rule out the possibility that the class of white things is a theory-laden category. Certainly with familiar words, common nouns have a privileged position and promote many more inferences than adjectives (Gelman & Coley, 1990; Gelman et al., 1986).

Although we have no evidence concerning whether children would attend to language form per se to determine whether a completely novel word maps onto a natural kind (e.g., to determine that a “fep” could refer to a member of a natural kind, but a “fep one” could not), we do know that children from 1½ years of age onward are
Input. Parental naming strategies (the context in which a word is used; the range of exemplars selected when the name is used) could provide indirect cues as to which names are natural kind terms. For example, the way superordinates are introduced suggests that they are less central to the identity of an object than are basic-level terms. In particular, parents “anchor” superordinates at the basic level (e.g., when teaching animal, parents are more apt to say, “This is a cat; it is an animal” than simply “This is an animal”) and generally fail to use superordinates to name just a single object (Blewitt, 1983; Callanan, 1989). Moreover, children’s assumptions about the input may combine with the variations in input (Callanan, 1989). In this way, different ways of speaking to children may do more than teach new vocabulary; they may be a means of marking which words map onto richly structured natural kinds.

Prototypes. Finally, the order in which category exemplars are introduced or assimilated could determine whether children treat a concept as richly structured. If a child first considers just the prototypical members of a category (e.g., for birds: robins, sparrows, and wrens), the category will be more coherent (see Mervis & Pani, 1980, for an excellent empirical demonstration of this phenomenon). Therefore, the child should be more likely to treat the category as a natural kind. In contrast, if the child is first introduced to all the anomalous and peripheral category members (e.g., parrots, hummingbirds, and flamingoes), the category may seem to lack sufficient coherence to be a natural kind. In this sense, prototypes may have functional utility: Children’s notorious difficulty in learning atypical category members (see Mervis & Rosch, 1981, for a review) could help keep categories coherent and so enable children to determine more quickly when they are natural kinds.

Summary

We hypothesize several components, acting in combination, that may contribute to children’s expectation that certain words map onto categories that are richly structured and inference-promoting. In the absence of conflict among them, the default is that the category name refers to a richly structured, inference-promoting category (natural kind assumption and linguistic transparency). However, this default can be overridden if the word fails to map onto a coherent category (coherence assumption). Children also gather additional clues from the linguistic form in which a word is embedded, the structure of the input that children hear, and the order in which category instances are introduced or assimilated (with children’s presumption of prototypes playing an important role). Although not all the evidence is in, preliminary results support the promise of this framework.

Summary and conclusions

In this chapter we have examined the interrelations between language, concepts, and the world in the developing child. By studying natural kind categories we have provided a test case of the more general question of how language structures thought. Language helps structure children’s categories, but the influence is subtle and works within the context of children’s nonlinguistic understanding.

We began the chapter by endorsing the claim that the everyday categories of adults, like the most sophisticated categories of scientists, are informed and held together by theories. This is a proposition that has recently received much support and is certainly not our own idea. From there, however, we went on to suggest six properties of theory-laden categories: They have a nonobvious basis, promote inductive inferences, are thought to embody an essence, incorporate anomalous members, observe division of linguistic labor, and are subject to revision. We argued that adults’ categories vary in the extent to which they embody these properties and accordingly vary in the degree to which they are theory-laden. Natural kind categories seem to be the best examples of theory-laden categories for adults, embodying all of these properties.

Next, we applied this framework to children’s categories, arguing that despite the fact that children lack an extensive knowledge base, their natural kind concepts exhibit many of the properties that characterize adults’ natural kinds. Our work, and the work of many others, leads to the conclusion that many of children’s categories promote inductive inferences, incorporate nonobvious properties, and reveal psychological essentialism. The jury is still out on the re-
The acquisition of natural kind terms

5. Unlike previous studies, this experiment tested children on properties that were already well known for prototypical category members. Thus, the question arises as to whether the children were making inductive or deductive inferences. We believe that children’s inferences can be characterized as inductive, because the atypical instances queried were novel (e.g., before entering the experiment the children did not know that dodo birds are birds or that pterodactyls are not birds), thus requiring inferences beyond what the children already knew with certainty. Furthermore, although the children themselves may have constructed more general arguments (e.g., “All birds live in nests”) from which deductive inferences could be drawn (e.g., “This [dodo bird] is a bird, so it too lives in a nest”), the construction of the general argument would itself have required an induction from the available evidence. See Gelman and Coley (1990) for more detailed discussion.

6. Note that although both perceptual learning and labeling may yield category change in the absence of theory, they can also induce more far-reaching theory change. That is, they may induce change not only in children’s beliefs as to which entities are category members, but also in more fundamental beliefs concerning what is central to the category.

7. We are not proposing that children possess a theory-independent sense of similarity or coherence. Rather, certain properties are assumed to be privileged from an early age (see R. Gelman, 1987; Medin & Wattenmaker, 1987).

8. Ontological constraints may play a large role in determining when to override linguistic transparency (see Keil, 1979).

References

Brown, R. W. (1956). Language and categories. Appendix to J. S. Bruner,


The acquisition of natural kind terms


