13 Essentialist beliefs in children: The acquisition of concepts and theories

Susan A. Gelman, John D. Coley, and Gail M. Gottfried

In their first few years of life, children are making sense of the world at two levels at once: at the fine-grained level of everyday object categories (deciding which things are trees and which are dogs and which are cookies), and at a broader level that some have called commonsense "theories." Both are remarkable achievements. First, consider categorization. If children's vocabulary is any indication, by the age of 6 they have carved up the world into thousands of distinct categories (Carey, 1978). Many children undergo a vocabulary "explosion" at roughly 18 months of age (Halliday, 1975; McShane, 1980; Nelson, 1973), when the rate of acquisition suddenly rises exponentially. One child studied in detail by Dromi (1987) produced as many as 44 new words in one week, and roughly 340 new words in her first 7 months of speech. No other species acquires symbolic communication at this rate. Even studies that successfully teach apes to acquire sizeable vocabularies in sign language are incomparable, with no noticeable vocabulary explosion (e.g., after more than 4 years of exposure to sign language, Washoe acquired only about 132 signs; Gardner & Gardner, 1989).

At around the same time that children learn to classify individual entities and undergo rapid vocabulary growth, they are developing broad systems of belief about the world. Not only do children learn to identify certain objects as "dogs," but they also learn that dogs belong to the class of animals, and that animals engage in characteristic biological processes such as growth, inheritance, and self-generated movement. Children are learning about physical laws such as gravity, mental states such as dreams, and social relationships within units such as families. They are learning where things come from, how things change over time, what causes an event to occur, and why. These belief systems include understanding causal relations, and they allow children to make predictions and provide explanations.

The writing of this paper was supported by NSF grant 91-00348 to S. Gelman. We thank Larry Hirschfield for his helpful comments on an earlier draft.
We believe that these two developments are not independent—that in fact, they go hand-in-hand. Children’s concepts are embedded in larger structures that resemble theories (see also Carey, 1985). In this chapter we illustrate the point with the domain of living things. There are two motivations for examining this domain. First, young children are fascinated by living things (especially animals) and devote considerable attention to them. Even infants pay special attention to things that move on their own (Poulin-Dubois & Shultz, 1990), and animal terms typically constitute a sizeable portion of children’s early vocabulary (Nelson, 1973). If we assume that children are most likely to build theories about objects of great interest, then the domain of living things is a likely place to look for theorylike beliefs. Second, living things figure prominently in debates concerning the domain specificity of children’s thought (see Keil, this volume; Carey & Spelke, this volume; Branon, this volume). It is unclear whether biology constitutes a separate domain for young children, and if so what form it takes. Thus, a close examination of children’s theories of living things may provide additional information on this issue.

The structure of the chapter is as follows. First we outline some properties of commonsense theories. Then we present a variety of studies suggesting that children appeal to something like theories when reasoning about living things. We advance the notion of “psychological essentialism” (Medin, 1989) to characterize the results of these studies, which examine children’s beliefs about causal mechanisms, innate potential, and maintenance of identity over transformations. Finally we consider alternative accounts of these data, and conclude that children appeal to theorylike entities that may derive from domain-general expectations.

What is a theory?

Everyday thought is theorylike, in the sense that people make use of unobservable, causal-explanatory constructs (see also Carey, 1985; Gopnik & Wellman, in press, for discussion). We distinguish between a commonsense or folk theory and a scientific theory (see also Brewer & Samarapungavan, 1991; Carey, 1985; Gopnik & Wellman, this volume; Kaiser, McCloskey, & Proffitt, 1986; Karmiloff-Smith & Inhelder, 1975; Keil, 1989; McCloskey, 1983; Murphy & Medin, 1985). A commonsense theory does not entail the detailed, explicit, formal understanding that Ph.D. biologists or physicists have. People’s beliefs seem to lack the coherence and systematicity typically associated with scientific theories (diSessa, 1988). Even when people have implicit understanding of a theoretical principle, they may be poor at stating that principle explicitly. And even when a hypothesis is strongly held, people seldom conduct rigorous or adequate experiments to test their hypotheses (Wason, 1960). In ordinary circumstances, people rarely engage in hypothetico-deductive reasoning (Shaklee, 1979).

Nonetheless, adult thought is theorylike in its appeal to domain-specific causal laws. For example, we believe that if one marble collides with another, it will cause the second marble to move; or that releasing a pebble from a height will cause it to fall. These mechanical cause/effect relations operate on physical objects only; mental states cannot be described in the same terms. Conversely, the laws of reinforcement apply only to entities capable of psychological processes (e.g., people, dogs), not to objects in the domain of physics. Biology, too, has its own laws relating to eating, breathing, inheritance, growth, and so on.

In order to account for these causal relations, people invent powerful, unobservable constructs. In the preceding physical examples, we appeal to force and energy to explain why one object causes another to move; we appeal to gravitational forces to explain why an object falls from a height. In psychology we refer to beliefs and desires. For living things we refer to genetic dispositions and innate potential. Force, gravity, beliefs, desires, and genes are not visible or easily measured. They are constructs designed to explain events that we can see.

People’s tendency to create explanatory constructs can lead us to classify together entities that have salient differences but share theory-relevant properties. For example, adults generally classify plants and animals together into a category of living things, largely because we believe that both plants and animals grow, need water, reproduce, and have the ability to heal themselves (Bakeman, Shatz, & Gelman, in press; Hickling & Gelman, 1992). Furthermore, adults sometimes refrain from classifying together things that seem superficially the same but differ in theory-relevant properties. In one such theory-based decision, adults exclude whales from the category of fish; although the two animals have similar appearances and habitats, there are important differences in underlying features such as breathing ability and blood temperature. Thus, theory-based classification is available to adults in our culture.

It is important to raise two additional issues that bear on theory-based categories. First, not all categories have this form—a point to which we return in the section, “How domain-specific are essences?” (cf. discussion in Markman, 1989). Second, classifications that appear theory-based are not always a direct consequence of theories. Specifically, people (especially children) may sometimes learn the “theory-based” classification before they learn the theory. For example, they may learn that whales are not fish before they learn that whales can’t breathe underwater, have warm blood, and bear live young. Learning that whales are not fish may even encourage a child to search for underlying differences between the two.

In this chapter we examine theorylike beliefs by discussing the unobservable constructs children invoke to account for biological properties and events. In particular, we focus on what Medin (1989; Medin & Ortony, 1989) calls “psychological essentialism.”
Essentialism

Psychological essentialism is the idea that people have an implicit assumption about the structure of the world and how it is represented in our categories. Specifically, people seem to assume that categories of things in the world have a true, underlying nature that imparts category identity. People's assumption has two parts: (1) that the world has a natural order that is independent of the observer (a realist assumption), and (2) that categories—and words referring to categories, such as common nouns—map onto that structure. On this view, categories are discovered rather than arbitrary or invented; they carve up nature at its joints. The underlying nature, or category essence, is thought to be the causal mechanism that results in those properties that we can see. For example, the essence of tigers causes them to grow as they do—to have stripes, large size, capacity to roar, and so forth. It is important to keep in mind that in discussing essentialism, we are endorsing a psychological claim about people's beliefs, not a metaphysical claim about the world (see also Medin & Ortony, 1989).

Essences are predictive of, but distinct from, more obvious or observable features (Gelman & Medin, 1993; but see Jones & Smith, 1993, for an alternative view). Surface features are correlated with and provide good clues about category membership (e.g., wings, feathers, and flight of a bird). Yet this information is fallible (not all birds have these features; some nonbirds do). In contrast, the essence may not be immediately observable, yet it is important because it determines category identity. It can be thought of as an unseen quality that is responsible for the observable features that hold together a category. Medin (1989) describes the essence as follows:

People act as if things (e.g., objects) have essences or underlying natures that make them the thing they are. Furthermore, the essence constrains or generates properties that may vary in their centrality. One of the things that theories do is to embody or provide causal linkages from deeper properties to more superficial or surface properties. (p. 1476)

Recently, several theorists have argued, independently, that an essentialist assumption is extremely common among adults. Atran (1990; this volume) reports that essentialist beliefs can be found across widely varying cultures. Mayr (1988) traced essentialist arguments back hundreds of years and proposed that a belief that an unchanging essence determines category membership was one of the biggest obstacles to grasping evolutionary theory. Although indirect evidence suggests that an essentialist stance is prevalent among adults in our culture, the idea has received surprisingly little direct confirmation.

Where does essentialism come from? Is it a basic human assumption, or simply a by-product of our culture and modern-day science? One way to investigate the issue is to study young children, who have not yet been exposed to formal scientific training. They are not taught about DNA or molecules in

nursery school, and parents rarely talk about the molecular and genetic distinctions between different species of plants or animals. Furthermore, years of experimentation have documented that children are especially attentive to surface appearances (Flavell, 1977). On many tasks children have difficulty looking beyond the most superficial cues (cf. Piaget's conservation task, in which children below age 6 or 7 fail to appreciate that the quantity of a liquid is unchanged as a result of pouring the liquid into a differently shaped container). If children assume the existence of nonobvious entities, and furthermore assume that such hidden entities can be critical to an object's identity, then we can infer that an essentialist assumption is powerful and basic.

There is already reason to suspect that children may be essentialists. To begin with, evidence is growing that children assume that events are caused (causal determinism; A. Brown, 1990; Bullock, Gelman, & Baillargeon, 1982; Gelman & Kalish, 1993). Studies with infants indicate that they seem surprised when viewing events with no apparent cause (Baillargeon, 1993). By preschool age, when children view an event that appears to violate known causal laws (e.g., a screen passing through a box), they attempt to dismantle the apparatus, apparently in search of a hidden causal mechanism (Chandler & Lalonde, in press). Furthermore, children have surprising difficulty grasping random phenomena. Piaget and Inhelder (1955) showed children random devices such as spinners, and found that the youngest children often insisted that they could predict where the pointer would fall. The same kind of pattern is found when children are asked to explain adverse events, such as someone falling ill. Children tend to blame the illness on the victim himself or herself, rather than allow that it happened randomly, an explanation known as "immanent justice" (Karniol, 1980; Kister & Patterson, 1980; Piaget, 1948; White, Elsom, & Prawat, 1978). In all these examples, causal determinism refers to events, but it could as easily refer to properties of objects. That is, children may reason causally not only about events such as falling ill, but also about object features, such as the spots on a giraffe or the ability of a rabbit to hop.

If children do assume causal determinism, then essentialism could be just one step away. That is, upon viewing events or features with no observable cause, children may impute internal, nonobvious, or invisible cause. This would be a way of resolving the apparent contradiction between no visible cause and the need for all events to be caused. Bringing in a hidden cause need not be a conscious or explicit strategy; rather, it may naturally follow from the assumption of causal determinism in such situations. Thus, it is at least plausible that essentialist reasoning is present even from earliest childhood.

What counts as evidence of essentialism?

People may believe that a category has an underlying essence even if they believe they will never know what it is. Using Medin's terminology,
they have an essence placeholder. We reason about categories as if they have an essence, even while the specifics of the essence are unknown or unknowable. One practical consequence is that it is difficult to find direct evidence of essentialism.

Simply forming a category is insufficient evidence for belief in an essence. At least in principle, one can form complex biological concepts without assuming essences. Pigeons appear able to classify trees or fish quite accurately on experimental tasks (Herrnstein, Loveland, & Cable, 1976), yet presumably pigeons don't assume essences. In a sense, then, the question is whether biological concepts (e.g., "tree," "dog") take a different form in young children versus pigeons.

We describe various kinds of evidence that are consistent with, and may provide some support for, psychological essentialism: (1) **Explicit articulation of an essentialist philosophy**. This type of evidence is the most direct, but least likely to be found – especially in young children, who are notoriously inarticulate when it comes to explaining their own reasoning processes. (2) **Appeal to invisible causal mechanisms**. An essentialist assumption may emerge as a reference to an unseen entity or quality that is intrinsic to an object, and causes surface features. For example, among educated adults in the United States, DNA is believed to be unseen (by the naked eye), intrinsic to an organism, and possessing the power to determine changes that occur via growth. (3) **Assumption of innate dispositions or potential**. People may refer to inborn capacities, presumably determined by the object's essential nature, that emerge later in life. For example, the human child is believed to have innate potential that is causally responsible for the development of crawling, walking, and language. (4) **Unalterability, or maintenance of identity over superficial transformations**. A belief in essences leads to a belief in unchanging category membership, even in the face of certain dramatic changes in observable properties (i.e., an assumption of unalterability; see Rothbart & Taylor, in press). For example, essentialism would support the belief that an animal should retain membership in a category despite growth, metamorphosis, or plastic surgery that greatly alters its appearance. (5) **Inductive potential**. If we assume that category members share an essence, then we can also assume that this essence leads to deep or unlimited commonalities among category members. For example, given an essentialist bias, giraffes would be assumed to be alike with respect to internal organs, skeletal structure, body temperature, means of nurturing their young, life expectancy, speed of locomotion, and so on. (6) **Taxonomies**. Atran (1990) proposes that evidence for essences can be found in consensus about the taxonomy within which categories fall, "despite obvious variation" among category members (pp. 6, 62). This may be the least direct of the various kinds of evidence, because of the assumptions it makes about taxonomies (which are debatable; see Collins & Loftus, 1975; Hampton, 1982) and their origins.

Combining any of the above pieces of evidence should provide even stronger evidence. For example, if one appeals to an invisible causal mechanism that is unalterable, produces unlimited commonalities (not just one effect), and commonalities that hold among all category members despite superficial differences, then we would have more compelling evidence for psychological essentialism.

In the following sections we review evidence for three of the more direct kinds of evidence for essentialism: appeal to invisible causal mechanisms, assumption of innate potential, and maintenance of identity over superficial transformations.

**Causal explanations**

A critical aspect of essences is their causal force. Locke (1894/1959) talks about the essence as the causal mechanism that gives rise to those properties that we can see. If children are essentialists, they should search for underlying causes that result in observable features (e.g., assuming some underlying nature that causes category members to be alike). "Features" include not only perceptual appearances, but also behaviors and/or events that are shared by category members. For example, the essence of a tiger causes it to have stripes, large size, capacity to roar, and the like. As mentioned earlier, there is some hint in the literature that children may assume that events are caused. There is also evidence that, when explaining events with no observable cause, children appeal to underlying causes (Shultz, 1982). For example, upon viewing a radiometer (a device that spins when light is beamed on it), children as young as 4 years of age typically said "yes" when asked if there was "some invisible thing that goes from the light to the propeller." How powerful is this assumption of underlying cause, and does it apply more for some domains than others?

We have conducted a study of children's causal explanations to examine whether children view underlying nature as causing natural events to occur (Gelman & Gottfried, 1993). Four-year-old children viewed brief videotapes in which actual animals and objects moved across a surface. All the items were unfamiliar, to ensure that responses did not simply reflect specific experiences with these particular objects, either direct observations or statements made by more knowledgeable others. After viewing each event, subjects were asked to make two judgments about causal mechanisms, an External Cause judgment ("Did a person make this move?") and an Internal Cause judgment ("Did something inside this make it move?"). In addition, children were asked to make an Immanent Cause judgment ("Did this move by itself?").

We also encouraged children to justify their responses after every question.

To determine whether children's judgments vary as a result of the ontological status of an object, the moving objects were chosen from three distinct categories: animals (e.g., chinchilla), wind-up toys (e.g., wind-up toy sushi), and transparent artifacts (e.g., plastic pepper mill). Children were randomly assigned
to one of two conditions that differed in the kinds of events presented. In the Baseline condition (also known as the Carried condition), all objects were pushed or carried by a person, whose hand was always visible on the videotape. Here, children should appeal to external cause, reporting that a person made all the objects move. In the second condition (referred to as the Alone condition), the objects started at a standstill and moved without any apparent human intervention. In this condition, we predicted that children should resist external explanations and more consistently appeal to inherent, internal causes. That is, because no person was visible and the object’s motion appeared to be self-generated, children were expected to impute an internal cause (Gelman & Kremer, 1991). This finding would provide evidence for a belief that an underlying essence caused the object to move.

Of interest is whether inferences about internal cause would be equally strong for the three kinds of objects. Past research has suggested that an assumption of internal cause should be powerful for animals (Massey & Gelman, 1988); it is unclear whether the same applies for artifacts, which typically move as the result of human agents. The transparent artifacts provide the strongest test of children’s belief that events without obvious external cause are interally or inherently caused. Because the children viewed transparent objects moving by themselves and could see that nothing was inside the objects, there was a potential conflict between the assumption of internal cause and the evidence provided by the children’s own eyes. Only if the internal cause assumption is very strong would it override the visual evidence.

Children’s responses to the causal mechanism questions (“Did a person make this move?” “Did something inside this make it move?”) were clearly different for the animals and artifacts. For the wind-up toys and transparent artifacts in the Baseline (Carried) condition, children were much more likely to attribute the cause of motion to a person than to anything inside. With animals, however, the pattern is reversed. Children regularly denied that a person made the animals move. This result is striking, given that the animals were carried and that the person’s hand was visible throughout the event.

When the object appeared to be moving alone (Alone condition), children were less likely to claim that a person caused the movement. Rather, they claimed that something inside the item made it move. Interestingly, children displayed this pattern not only for the animals but also for the artifacts, including the transparent ones. That many children attributed internal mechanisms to transparent objects with no apparent insides shows the strength of their belief that internal mechanisms cause objects to move on their own.

Children’s responses to the immanent cause question showed the same pattern as the internal cause question, only more dramatically. Over 90% of the children in each condition claimed that the animals moved by themselves. Even when a person bodily carried the animal from one end of the screen to the other, children insisted that the animal itself was responsible (e.g., one child responded for each of the animals, “It wasn’t [intended] for moving with a person”). For toys and transparent artifacts, only children in the Alone condition said the objects moved by themselves. The absence of an external agent led them to endorse immanent cause.

It is somewhat misleading to look only at children’s yes–no responses to the question, “Did it move by itself?” When children responded that an object made by itself, they could simply have been reporting their observations, rather than implying any particular cause. In other words, “It moved by itself” could be equivalent to claiming that a person did not make it move. However, an analysis of children’s explanations for how the objects moved by themselves shows that children provided causal mechanisms that were distinct to the domain of the object.

The children seemed to have particular difficulty explaining the artifact events. On average, children said that they just didn’t know how it moved or that some supernatural agent was involved, about 5 times more often for artifacts (toy or transparent) than for animals. Of the more interpretable responses, children primarily used two kinds of justifications not found with animals. First, they often mentioned a person as an agent. Even though they saw on the videotape that the item started from a standstill and moved on its own, they mentioned some person winding up, pushing, or moving the object. The second response was to refer to other natural causes that were not inherent to the object per se, such as electricity, batteries, or magnets. Children even occasionally tried to reconcile their judgments with what they saw on the video: One child suggested there were “invisible batteries” making a transparent artifact move, and one suggested that an “invisible person” made a wind-up toy move.

For the animals, children’s justifications were nearly always coherent. The children never referred to an external agent, as they did with the artifacts. Rather, 93% of the justifications referred to properties intrinsic to the animal, such as its bones or other parts. The children also occasionally claimed that the animal caused itself to move, with justifications such as, “Because he makes it move himself” or “Only itself can do it.” Although children claim that all the objects move by themselves, how they resolve the question of causal mechanism differs sharply by ontological category.

The results of this study show that when children cannot find an external cause, they invoke one that is internal or immanent. Interestingly, children invoke internal cause for animals and equally for both opaque and transparent artifacts, but they consistently deny any external cause to explain the biological events, even when a human bodily carries the animal. Rather, biological events are viewed as resulting from immanent cause. Children regularly appeal to intrinsic factors even without knowing the internal mechanism (e.g., responding “it just did it itself”). Thus, children seem to be displaying a belief in some underlying causal mechanism, without knowing exactly what that mechanism entails.
Innate potential

A second possible indication of essentialism is belief in innate potential, that a set of characteristics will unfold with maturation, even though they show no sign at birth. For example, a lion cub has the potential to grow into something large and fierce, even though when it is born it is small and helpless. The fact that such characteristic attributes emerge so predictably suggests that the individual possesses nonobvious, intrinsic qualities. To explain developmental changes like this, we as adults might say that lions have an essential nature that is responsible for how they grow.

To test whether preschool children have an idea of innate potential, Gelman and Wellman (1991) conducted a study that can be thought of as pitting nature against nurture. On each of a series of items, 4-year-old children learned about a baby animal that was raised among members of a different species, in an environment more suited to the adoptive species. We attempted to describe the environment as an interactive, nurturing one, in order to increase the plausibility that the environment could exert important effects. Thus, in each scenario, the adopting species were said to “take care of” the adopted animal. Also, every picture of the adoptive species contained a mixed-age, familylike grouping. The main issue was what children predicted about the animal’s appearance and behavior, after it reached maturity. Would it show the potentialities inherent in its category membership, or would it display the properties one would expect from its environment?

For example, children first saw a picture of a baby kangaroo that looked like a nondescript little blob of an animal. They learned that the baby kangaroo was taken to a goat farm when she was a baby, and raised by goats. Then they saw a picture of the goat farm (4 adult goats and 2 kids). We stressed that the baby grew up without ever seeing another kangaroo.

Both pictures – of the baby and of the contrasting environment – remained in view while children were asked two questions about how the animal would be after it grew up. In this case they heard, (1) Was she good at hopping or good at climbing? (2) Did she have a pouch or was she without a pouch? We pretended all the properties, to make sure that children of this age could answer them correctly when there was no conflict – for example, they said that goats are good at climbing and don’t have a pouch.

The results are as follows. Children nearly always answered on the basis of category membership or innate potential. For example, they said that a baby kangaroo raised among goats will grow up to hop and have a pouch. Children were not just reporting associations to the category label (for example, saying “pouch” as an unthinking response to “kangaroo”), because they were sensitive to question type. Children relied more on innate potential when the question concerned a behavior (for example, what noise a cow will make) than when it concerned a static property (for example, whether a cow will have a straight tail). If children were simply reporting associations with the category label, there should have been no difference between behaviors and properties.

However, we wanted to obtain more direct evidence that children were not simply reporting category associations. We conducted two additional studies on the issue. In one study, we asked children about the properties of baby animals. For example, we asked the children whether or not the kangaroo, when it was a baby, had a pouch or could hop. If children were simply reporting category associates, they should report that the baby kangaroo has all the properties of the adult kangaroo. However, if children were considering the particulars of the individual animal, then they should say that the baby kangaroo does not yet possess its full adult properties. In addition, we included inherent properties (e.g., did it have eyes?) and impossible properties (e.g., did it have wings?) as baseline control questions.

Results showed that the children attributed the behavioral properties to the babies as often as the intrinsic properties, suggesting a possible tendency to report category associates for questions about behavior. However, children generally reported that the babies did not have the physical features of the adult animal (e.g., the pouch). Thus, children do not simply report category-associated properties upon hearing questions about category members. Furthermore, the data converge to suggest a stronger understanding of innate potential: Animals will develop properties that they don’t yet possess as infants, despite being raised in an unusual environment with no same-species cohort.

Gelman and Wellman (1991) also conducted a “nature-nurture” study with seeds. For example, children saw a seed that came from an apple and was planted in a flowerpot. The test question was, “When that seed grew, what popped up out of the ground? Was it an apple tree or a flower?” Children saw pictures of the seed and of the environment (e.g., apple seed, flowerpot). Seeds provide a strong control for children’s tendency to report category associates, because specifying the origin of a seed does not entail stating its category identity (e.g., “comes from an apple” differs from its original identity of “apple seed” and its future identity of “apple tree”). In addition, a seed looks nothing like its eventual endstate (plant or tree). Finally, the use of seeds allows us to examine a very different kind of parent-offspring relationship, in which characteristics cannot be transmitted by means of modeling, reinforcement, or training. Results of this study showed that 58% of the younger 4-year-olds and 92% of the older 4-year-olds answered primarily on the basis of innate potential. The mixed performance of the young 4-year-olds appears to be due to a less-developed understanding of the relation between seeds and plants at that age (Hickling & Gelman, 1992). Nonetheless, it is striking that most of the children consistently reported that a seed has the innate potential to develop in accordance with the parent species. Four-year-old children act like essentialists. They assume that members of a category share an innate potential, and that innate potential can overcome a powerful environment.
Maintenance of identity

An underlying essence would allow individuals to undergo marked change yet retain their identity. We know that adults in our culture believe that radical changes, such as metamorphosis, are possible (Rips, 1989). Furthermore, Keil (1989) has shown that second graders (though not preschoolers) realize that animals but not artifacts can maintain identity over such transformations. Children were shown pictures of animals, then told about transformations performed by doctors that changed the characteristic features of the animal into those of another animal. For example, a tiger had its fur bleached and a mane sewed on, so that it now resembled a lion. Children were then asked whether the posttransformation animal was a lion or a tiger. Second and fourth graders maintained that the identity would not change. This finding implies an early-developing belief that animals, but not artifacts, possess essences that are responsible for maintenance of identity.

Work with younger children demonstrates a similar kind of understanding. Gelman and Wellman (1991) used a paradigm very similar to that of Keil (1989), but with simpler transformations: Each item had either its “insides” or its “outsides” removed. Test items were selected to be clear-cut examples (for adults) of objects for which insides, but not outsides, are essential. For example, blood is more important than fur to a dog; the engine of a car is more important than the paint. As a control, we also selected a set of items for which the insides are not integral parts (e.g., a jar, a refrigerator).

Gelman and Wellman asked 4- and 5-year-old children to consider three transformations: (1) removal of insides (e.g., “What if you take out the stuff inside of the dog, you know, the blood and bones and things like that, and got rid of it and all you have left are the outsides?”); (2) removal of outsides (e.g., “What if you take off the stuff outside of the dog, you know, the fur, and got rid of it and all you have left are the insides?”); (3) movement (e.g., “What if the dog stands up?”) as a control. For each transformation, children heard two questions: (1) identity (“Is it still a dog?”), and (2) function (“Can it still bark and eat dog food?”).

The results from every test item indicate that children say that if you remove the insides, the identity and function of an object will change. Not so for the outsides, even when removing them would sharply change the object’s appearance. As predicted, the children correctly reported that the identity of the containers (e.g., refrigerator) would not change if the insides were removed.

An additional series of studies by Rosengren, Gelman, Kalish, and McCormick (1991) examines children’s understanding of maintenance of identity, using the natural biological transformation of growth. Rosengren et al. reasoned that an important piece that may be missing from past research is what mechanism is underlying the change. In other words, children may be sensitive to whether the mechanism is a natural biological transformation or one that defies biological laws. The implication is that, even though children report that some transformations lead to identity change, they may realize that natural transformations (such as growth) do not.

Rosengren et al. conducted a series of experiments demonstrating that children as young as 3 years of age expect animals to undergo changes over time (via growth) without affecting identity, that children believe that such changes are strongly constrained (e.g., one can only get bigger not smaller over time), and that these changes are specific to the domain of living things. For example, 3-year-olds, 5-year-olds, and adults were shown a picture of an animal and told, “Here is a picture of Sally when Sally was a baby. Now Sally is an adult.” They were then shown two pictures: one identical to the original and the one the same but larger, and were asked which was a picture of Sally “as an adult.” At all age groups, subjects tended to choose the larger picture, showing that they expected the object to undergo change in size with growth.

By 5 years of age, children realize that growth is inevitable. For example, in another condition children saw a picture of a juvenile of a species that undergoes radical metamorphosis (such as a caterpillar). They then saw a picture of the same creature, only smaller (e.g., a smaller caterpillar), and a picture of a larger animal differing in shape (e.g., a moth). Again, subjects were asked to choose which picture represented the animal after it became an adult. Three-year-olds were at chance, but 5-year-olds chose the metamorphosed animal significantly above chance levels. By the age of 5 years, then, children believe that an individual can naturally undergo even substantial shape changes over time.

These data are demonstrational. The studies conducted by Gelman and Wellman (1991) and by Keil (1989) show that children realize that sometimes, the features most critical to an object’s identity may be internal and nonobvious. The experiments of Rosengren et al. (1991) demonstrate maintenance of identity over changes wrought by growth. In both cases, children endorse the possibility that objects have important underlying properties.

Other evidence

A variety of additional studies also provide evidence consistent with an essentialist bias. On these tasks, subjects reveal that they look beyond surface similarity when reasoning about categories. Two and one-half-year-old children appreciate that animal categories support inductive inferences regarding familiar properties (Gelman & Coley, 1990), and 3- and 4-year-olds use categories (animals, plants, substances, and artifacts) to guide inferences about novel properties (Gelman, 1988; Gelman & Markman, 1986, 1987). Four-year-old children also appreciate the importance of internal, intrinsic causal mechanisms for living things and artifacts. For example, children report that a bird flies because of its heart and muscles, that a car moves uphill because of its motor, and that a flower blooms on its own (Gelman & Kremer, 1991; see also Gelman, 1990). Children also realize that human intervention
Alternative explanations

The studies reviewed in this chapter do not provide direct evidence for psychological essentialism. They do, however, strongly suggest that children appeal to invisible causal mechanisms, assume innate potential, maintain identity over superficial transformations, and assume unlimited commonalities among category members when they reason about objects. Although not constituting unambiguous evidence for essentialism, these studies do allow us to rule out several competing explanations.

The first, most obvious point is that children’s judgments do not simply reflect reliance on perceptual similarity. For example, maintenance of identity over superficial transformations requires overlooking appearances (e.g., perceptual similarity alone would not predict that a caterpillar can become a butterfly). Although appearances are important, preschoolers can go beyond them.

Second, children are not simply reporting observations from past experience. It is unlikely that children can perceive essences, because many biologists and philosophers insist that animal species do not have essences (Mayer, 1988). Furthermore, experience alone does not present inheritance mechanisms or immanent causation as observable entities. Experience also would not be sufficient to account for the systematic errors that children make (cf. Taylor & Gelman, 1991, in which children reported that gender-role properties are innately determined).

Finally, children are not explicitly taught an essentialist philosophy. In fact, considering the input, it is remarkable that children hold any consistent set of beliefs. Children hear stories in which princess turn into frogs, people become statues, and elephants hatch out of bird eggs. Even in ordinary language, biological concepts—concerning disease, growth, or genetic relationships—leak over into other domains (computer viruses, sick jokes, dying cars, grandfather clocks, fabrics that breathe, a growing national debt). Still, it may be that these ideas are marked in some way as figurative or fictional. To provide a fair look at the issue, we need to see how parents discuss biological concepts in ordinary speech.

We recently gathered data on parental input to 20- and 36-month-olds (Gelman, Coley, Rosengren, Hartman, & Pappas, 1993). Thirty-two mothers and their children came into our laboratory and looked through two books together, one with scenes of animals and the other with scenes of artifacts. We created the books so that they would be as similar as possible in every way except for the objects shown on each page. Every page included objects in which appearances were potentially misleading, as we thought it might prompt parents to talk about nonobvious ways that categories are constructed. For example, on one page parents saw a scene with two bats, a blackbird, and a cave.

We videotaped the sessions and then analyzed the verbal transcripts and both parent and child gestures for any mention or reference to categories. We were especially interested in talk about anything that might be construed as an essence.

We found that, although parents placed great emphasis on categories, frequently pointing out category relations and similarity among category members, they rarely mentioned nonobvious features such as insides, innate potential, origins, or essences. For example, only one mother of the 20-month-olds (out of 16) discussed insides at all, and she made only one mention; origins were never discussed in this age group. Among mothers of the 36-month-olds, insides were mentioned a total of 9 times, with 8 of the 9 mentions referring to the insides of artifacts. In other words, on average only 10 out of 4,372 codable utterances (0.2%) concerned insides. Origins were mentioned a total of 3 times, meaning that on average only about 0.07% of codable utterances yielded discussion of origins. Thus, overt discussion of biological essences was vanishingly rare, even in a sample that was highly educated and motivated.

It seems, then, that children’s belief in the importance of nonobvious properties cannot be attributed to what they are directly taught by parents. Neither can it be “read off” their experiences of the physical world, nor reduced to a general similarity bias. Essentialism is not directly available in either the physical environment or the parental input.

How domain-specific are essences?

An essentialist stance appears to describe a variety of concepts, not only within the biological realm (as illustrated earlier; see also Mayer, 1991; Atran, 1990), but also including the following domains (although the authors cited do not always interpret their findings as evidence of essentialism): beliefs about ethnicity (Rothbart & Taylor, in press; Stoler, 1992, in press), personality traits (Yuill, 1992); race and occupation (Hirschfeld, this volume); gender (Gelman, Collman, & MacCoby, 1986; Taylor & Gelman, 1991); the workings of the mind (Wellman, 1990); social attributions (Hilton, 1992); and perhaps physical causality (Shultz, 1982). We illustrate this point with traits.

Yuill (1992) proposes that a trait such as generosity does not merely describe a cluster of related behaviors (sharing, helping, etc.), but rather is an inner quality that causes one to demonstrate these behaviors. The link between an overt behavior and the trait that explains it is analogous to the link...
between a biological structure or process and the essence that causes it (Gelman, 1992). Just as the trait of shyness can cause a person to avoid large parties, so does the essence of a panda cause it to have black-and-white fur and to eat bamboo. The person and the panda are each hypothesized to have an underlying quality (shyness or panda genes, respectively) that gives rise to certain observable properties as well as other, less obvious ones (beliefs and desires in one case; biological structures and processes in the other).

Confronted with this variety of essentialist thinking, some scholars have proposed that an essentialist assumption is domain-general (Medin, 1989). Mayr (1991: 41) implies a similar position:

Essentialism’s influence was great in part because its principle is anchored in our language, in our use of a single noun in the singular to designate highly variable phenomena of our environment, such as mountain, home, water, horse, or honesty. Even though there is great variety in kinds of mountain and kinds of home, and even though the kinds do not stand in direct relation to one another (as do the members of a species), the simple noun defines the class of objects.

Yet proposing an entirely domain-general essentialist assumption is on the face of it somewhat problematic: What is the evidence for adherence to underlying causal properties for wastebaskets? Indeed, Keil (1989) provides ample evidence that natural kinds are more likely than artifacts to be treated as having essences, even by the second grade. Similarly, Gelman (1988; Gelman & O’Reilly, 1988) finds that the inductive potential of animal categories is greater than that of artifact categories, suggesting that animal kinds are assumed to have a more tightly knit causal structure. The results from the Gelman and Gottfried study (1993) also revealed distinct differences in children’s explanations of animate versus inanimate motion, even when the source of visible movement was equated. Thus, a wholly domain-general essentialist assumption seems problematic.

Specifying the exact nature of the proposed essentialist bias requires explaining this pattern of broad but not boundless applicability. There are at least four potential classes of explanation: (1) borrowing from a base domain, (2) domain specificity in a broad, undifferentiated domain, (3) multiple domain-specific notions, and (4) domain generality with different, domain-specific instantiations. We favor the last possibility, although all four are worthy of serious consideration.

**Borrowing from a base domain**

First, it may be that essentialism begins as a domain-specific assumption, which is then analogized to other domains. This possibility requires that there is a base domain to which essentialism applies most readily or naturally. Biology is one candidate for such a basic essentialist domain (Atran, 1990), from which the essentialist stance may be extended, by analogy, to other domains. This model is what Hirschfeld (this volume) refers to as the “naturalization” model (see also Atran, Boyer, this volume, for supporting arguments).

Whether biology is the most plausible base domain is an open question. The evidence we have detailed here suggests that biological kinds appear to have essences for very young children. Additionally, biological essences appear to be universally assumed (Atran, 1990), whereas other sorts of essences, such as traits, may not be (Yuill, 1992). However, the domain of social interactions may also form the base. Dwyer (1976: 433) suggests that, for the Rosai (a group of people living in the Papua New Guinea Highlands), “species . . . in my interpretation, share an essence which is abstracted from human social structure . . . . This essence approximates symbolically, though it is not standardly labeled as such, a lineage (or kin descent group).”

One advantage to the naturalization model is that, because analogies are imperfect, it predicts some variation in essentialism is instantiated. Indeed, some important differences do emerge between different views of essences. For example, psychological traits can be altered; biological essences are immutable. Essences always have some outward manifestation (no matter how subtle); traits apparently need not. Essences hold true for all members of a category and so concern category identity, whereas traits distinguish people from one another and so concern individual identity. The naturalization model handles this variation by positing that traits only resemble essences—they are not equivalent.

A problem we have with this view is that it would require the analogizing to take place extremely early in development. Even preschool children have many kinds of categories that appear to have essences, including biological, psychological, and social categories. For example, Taylor and Gelman (1991) found that children are more likely than adults to apply an essentialist model to categories of gender (e.g., boy, girl). If essentialism is generalized so rapidly and widely with so little instruction, it calls into question what a domain is, and how useful or appropriate a domain-specific approach is.

**Domain specificity within a broad domain**

A second possible approach is to posit that essentialism is domain-specific, but that early in development children’s domains are relatively undifferentiated and broader than those of adults. For example, children may have an undifferentiated domain of social things, in which are lumped together things that for the adult will be differentiated into biology, psychology, and social interaction (Carey, 1985). Or, children may have an undifferentiated domain of biological things, in which sex, occupation, race, caste, and ethnicity are lumped? That is, distinctions that for the adult are eventually viewed as social constructions could be believed to be biological categories by children. The plausibility of this position rests in part on how broadly children assume.
essences. If the assumption extends to physical causality as well as psychological and social causality, then the "broad domain" position is less compelling. The plausibility also rests on the contention that children fail to differentiate among biology, psychology, and social convention, a matter of current debate (Carey, 1985; Inagaki & Hatano, 1988).

**Multiple domain-specific essences**

Another way to resolve the issue of broad but not unbounded essentialism is to posit multiple domain-specific notions of essentialism. On this view, different kinds of categories (living kinds, social kinds) are independently imbued with their own kind of essence. This explanation receives some compelling support (see Hirschfeld, this volume), although its plausibility would be limited if too many distinct kinds of essences were discovered. That is, it would be unparsimonious to propose one kind of essence to explain children's appeal to invisible causal forces such as wishes, thoughts, and desires as demonstrated in theory of mind research, another kind of essence to explain children's appeal to invisible causal forces in physics, a third to explain children's appeal to invisible causal forces in biology, and a fourth to explain children's appeal to invisible causal forces in the social realm. The plausibility of this approach may rest in its ability to incorporate these varied findings.

**Domain generality with domain-specific instantiations**

A final possibility is that essentialism is a domain-general assumption, but one that gets invoked differently in different domains, responding to the causal structure of each domain. On this view, the assumption of essences is a general one and so readily applies to new domains that have never before been encountered (e.g., viruses, computers). However, it does not sensibly apply to all domains. The analogy we have in mind is that an essentialist assumption is like a hammer. A hammer is not specifically restricted for use with nails, and can be used with a variety of other objects as well (e.g., screws, staples). Nonetheless, the hammer is designed to fit best with nails or naillike objects, and cannot be used to drive all objects (including books and hats) into walls or boards. Similarly, an essentialist assumption may be "designed" in such a way that it functions only when it meets domains with the appropriate features.

On this fourth view, essentialism would have to begin at a level of generality sufficient to encompass beliefs about traits, biological essences, beliefs and desires, race and gender, and any other categories that children treat as having essences. Children's assumptions may be something like causal determinism coupled with a willingness to consider causes inherent to the object. These assumptions would yield different implications in different domains, given the information provided in the environment. In the case of an animal, the child would notice the animal moving on its own, would see no apparent external cause (either human or mechanical), and so would conclude that some inner, inherent nature is responsible for its movement. R. Gelman (1987, 1990) proposed a similar principle, an "innards principle," to account for her data regarding movement judgments and descriptions of insides that children report for animals versus inanimate objects (Massey & Gelman, 1988). In contrast, in the case of a wastebasket, any behaviors or functions of the object could be readily traced to the people who made and use the wastebasket; hence, there would be no need to appeal to properties inherent in the object or a wastebasket essence.

We favor this last account regarding the domain generality of essentialism for several reasons: It predicts a broad yet not promiscuous application of essentialism, it allows essentialism to emerge in even novel domains, it is consistent with the developmental evidence that essentialism emerges early, and it provides a parsimonious account of why essentialism emerges in multiple domains. However, at present it remains speculative.

The discussion of domain generality raises questions concerning the relation between essentialism and commonsense theories, to which we turn now.

**Are essences components of theories?**

Essences look like they could be a component of a theory: They are unobserved, imputed entities that are assumed to have causal force. On the surface, then, it would seem that evidence for essentialism could constitute evidence for a theory. In other words, if children have essentialist beliefs about categories of living things, this could imply that children have a theory of living things.

On the other hand, theories are by their very nature domain-specific, and we have just argued in the previous section that essentialism may be a domain-general bias. Moreover, regardless of the stand one takes on the domain specificity of essences, essentialism does not in principle require a theory. It could be an isolated assumption about category structure that is independent of larger belief systems (Atran, this volume). Theories may follow after the fact, and indeed there is some evidence that larger belief systems show much more variability and malleability than essentialism. Furthermore, there is evidence that at least part of the time, it is the adults who form theories, encode them in language, and then pass them down to children. Children assume that there is much more in common to members of a category than meets the eye, but do not put forth detailed hypotheses about the biological processes involved.

We believe that an essentialist assumption fosters theory building and eventually becomes part of a number of distinct theories, even if it is not part of a theory from the start. Developmentally, there may not be a qualitative break between where an essentialist assumption ends and a theory begins. It
is again useful to invoke the distinction between scientific beliefs and commonsense beliefs. The question is not whether children have scientifically accurate classifications, nor whether they hold logically consistent belief systems. Rather, the question is whether children reason about causality, make ontological commitments, and hold interrelated beliefs. An essentialist assumption fosters looking for causes, expecting inductions within categories, nonobvious similarities, and so on. Even if these beliefs are not (at first) tightly structured into a coherent, domain-specific system, they are the building blocks or components of later theory building.

Conclusions

We have suggested in this chapter that young children seem to carry with them an essentialist bias, a bias that is then extended into adulthood. Even if we don’t carve up the world at its joints, an essentialist assumption makes it seem as if we do. This belief about the structure of the world is both adaptive and pernicious. Psychological essentialism is adaptive in that it leads us to search for new knowledge and to revise categories on the basis of that knowledge. But this kind of belief is potentially dangerous as well: It may promote stereotyping and inflexible thinking. As Gould (1990: 73) notes: “Some classifications channel our thinking into fruitful directions because they properly capture the causes of order; others lead us to tragic and vicious errors (the older taxonomies of human races, for example) because they sink their roots in prejudice and mayhem.”

But how pervasive (i.e., domain-general) is an essentialist assumption? Experimental evidence for an essentialist world-view is at present most abundant in the biological domain. That is, when reasoning about animals, preschool children expect category members to maintain their identity over superficial transformations, to have innate potential, to have rich inductive potential, to be driven by immanent causes. Nonetheless, children appear to have something like an essentialist stance in a variety of other domains as well. Accordingly, we have proposed that essentialism in children stems from a domain-general assumption. To state it rather crudely, children assume that events and features are caused, and appear biased to search for internal or inherent causes. These general strategies combine to encourage people to construct essentialist accounts when encountering events or features with no external causal mechanism. This would include a wide variety of seemingly natural behaviors or feature clusters, ranging from the flight of a bird to the customs of an ethnic group.

Beyond what it tells us about essentialism, this work suggests that early understanding of living things appears to be theorylike. An essentialist assumption, like a theory, can foster a search for invisible causal mechanisms responsible for an object’s actions, development, and maintenance of identity. An essentialist assumption, like a theory, leads to the expectation of nonobvious similarities among category members, and unlimited inductive potential. An essentialist assumption does not constitute a theory per se, but it does contain several key elements of a folk theory.

To conclude, concept acquisition and theory building seem to go hand-in-hand from the earliest points in development. Cognitive development is not a process of building up from simple units (such as concepts) to larger structures (such as theories). Both are developing simultaneously. We are unlikely ever to find a point when children have concepts without theorylike beliefs.

Notes

1. Note, however, that it is unclear that individual scientists’ theories are as objective, consistent, etc., as is often assumed (Brewer & Samaranopagan, 1991).
2. Historically, there are many different essentialist beliefs, entailing distinct kinds of claims (Schwartz, 1977). Here we attempt to characterize some components of everyday, lay essentialist assumptions. We realize that these will not necessarily overlap with essentialist beliefs of any one philosopher or scientist.
3. We moved the transparent artifacts by means of an attached clear thread that was not visible on the videotape.
4. We thank Grant Gutheil for suggesting this experiment.
5. It may also imply that certain things that adults treat as biological are omitted, because for adults a complete domain of biological things might include viruses, bacteria, and plants.

References

Essentialist beliefs in children


Essentialist beliefs in children


