

# Discerning the Division of Cognitive Labor: An Emerging Understanding of How Knowledge Is Clustered in Other Minds

Frank C. Keil<sup>a</sup>, Courtney Stein<sup>b</sup>, Lisa Webb<sup>a</sup>, Van Dyke Billings<sup>c</sup>,  
Leonid Rozenblit<sup>a</sup>

<sup>a</sup>*Department of Psychology, Yale University*

<sup>b</sup>*Department of Psychological & Brain Sciences, Dartmouth College*

<sup>c</sup>*New Haven Public Schools*

Received 5 October 2006; received in revised form 25 February 2007; accepted 2 March 2007

---

## Abstract

The division of cognitive labor is fundamental to all cultures. Adults have a strong sense of how knowledge is clustered in the world around them and use that sense to access additional information, defer to relevant experts, and ground their own incomplete understandings. One prominent way of clustering knowledge is by disciplines similar to those that comprise the natural and social sciences. Seven studies explored an emerging sense of these discipline-based ways of clustering of knowledge. Even 5-year-olds could cluster knowledge in a manner roughly corresponding to the departments of natural and social sciences in a university, doing so without any explicit awareness of those academic disciplines. But this awareness is fragile early on and competes with other ways of clustering knowledge. Over the next few years, children come to see discipline-based clusters as having a privileged status, one that may be linked to increasingly sophisticated assumptions about essences for natural kinds. Possible mechanisms for this developmental shift are examined.

*Keywords:* Cognition; Cognitive development; Conceptual change; Concepts; Social epistemology; Philosophy of science; Deference; Distributed cognition

---

## 1. Introduction

As adults we all believe that knowledge is not distributed smoothly and homogeneously in the minds of others. Instead, we assume that bits of knowledge and understanding cluster together in ways that reflect different areas of expertise. This sense of different areas of expertise not only helps us to seek out further information or advice, but it also enables us to

---

Correspondence should be addressed to Frank C. Keil, Yale University, Department of Psychology, P.O. Box 208205, New Haven, CT 06520-8205. E-mail: frank.keil@yale.edu

evaluate the likely quality of knowledge in other minds. If an individual offers an explanation or advice that is not within his normally considered realm of expertise, we are inclined to be more skeptical about its value (Fricker, 2002; Ward, 2006).

The distribution of different clusters of knowledge in other minds is a fundamental aspect of culture and acts as an overlay on the more common forms of knowledge that spread throughout a culture in patterns of contagion governed by cognitive and cultural constraints (Sperber, 1994, 1996). The distribution of knowledge is also influenced by the inevitable divisions of labor in a culture, which have corresponding divisions of cognitive labor. At some level, members of a culture must have some sense of the division of cognitive labor if they are to build their knowledge on the expertise of others.

For centuries, scholars have examined the divisions of labor that occur in cultures and the extent to which they are inevitable outcomes of culture and economic systems (e.g., Durkheim, 1893/1997; Smith, 1776/1904), but there has been far less discussion of the division of cognitive labor and how people construe those divisions. The ability to master these divisions and to use them to advance one's own knowledge is considered an essential part of science (Kitcher, 1993, 2001) and of our ability to evaluate the quality of information that we acquire through social transmission (Goldman, 2002). This article explores the development of the ability to understand what others know, with a particular focus on what we will call "disciplinary knowledge"; that is, knowledge clusters of the sort that are at the core of scientific disciplines.

There are several distinct ways to think of the distribution of knowledge in other minds, all of which correspond to real-world patterns and which could be used to chart different forms of expertise. Four of the more common ways are based on goals, access, objects, and disciplines.

Goal-based clusters arise from acquiring knowledge that enables one to achieve a goal; for example, if one has the goal of winning long-distance solo sailing races, one is likely to have a greater than average knowledge of the stiffness and strength of carbon fiber composites, offshore weather patterns, the effects of sleep deprivation, and fluid dynamics. Higher levels of knowledge in each of these areas converge to support the goal of solo sailing racing. Just as ad hoc categories can be organized around goals (Barsalou, 1983), so can clusters of knowledge.

Access-based clusters arise from a person's situation in life. Affluent people might have greater than average knowledge of fine wines, sports cars, and ski resorts, whereas homeless people might have greater than average knowledge of policing practices, outdoor temperatures, and soup kitchens. One's proximity or ease of access to various resources is the prime determinant of these knowledge clusters.

Object-based clusters can correspond to largely associative knowledge of the sort normally attributed to trivia buffs, fans, or fanatical collectors. Thus, an Edsel fanatic might know more than average about everything associated with that ill-fated car model from the 1950s such as its list price, its namesake in the Ford family, its economic demise, and references to Edsels in later literature. If a category is at a low enough level one can attempt to learn everything and anything associated with members of that category. (Goal-, access-, and object-based clusters can interact as well. For example, some goals can influence access, and some kinds of access can influence areas of trivia knowledge.)

Finally, discipline-based clusters, the focus of this article, are clusters of knowledge based on a set of deeper principles, often of a causal nature. We can understand the coherence of a

knowledge cluster in which an individual knows why spinning tops stay up, why pendulums swing at certain rates, and why billiard balls bounce as they do. That domain corresponds to an expertise in physical mechanics and is highly generative in predicting a potentially unbounded set of other phenomena to which that expertise would also extend.

Discipline-based clusters are an essential part of how we understand knowledge in the natural sciences and, by many accounts, the social sciences as well. As adults, we assume that there are stable causal and relational patterns in the world that, when partially understood, allow one to have a greater than average understanding of a wide range of phenomena arising from those patterns. Someone who understands the patterns involved in physical mechanics will have greater than average insight into spinning tops, projectile motion, levers, systems of pulleys, and any other phenomena in which those patterns are central. Someone who understands the patterns involved in gene regulation and expression will have a greater than average understanding of embryology, inheritance of traits, cancer, metamorphoses, and any other phenomena in which those patterns are central. Thus, one way to infer how knowledge clusters in other minds is to assume that someone who can answer how and why questions about a phenomenon is likely to understand other phenomena that arise from the same causal and relational patterns, even if they involve dramatically different surface objects. Such an inference does require some hunches about phenomena that are in the same domains of causal regularities or other underlying patterns. In this article we explore the emergence of this ability to infer discipline-like areas of expertise from real-world regularities, assuming that the child's grasp of causal and relational patterns will be related to the degree to which that child has intuitions about discipline-like areas of expertise.

Not all academic disciplines cohere because of a set of causal regularities that set them apart (Hacking, 1999). Most mathematicians would consider their field as cohering around a set of abstract formal principles that help explain, in an acausal manner, the structural properties of mathematics. In the humanities, expertise often coheres in departments in ways that seem more object or goal centered. Members of an English department might consider their expertise to cohere around information associated with bodies of text. Attempts to understand English literature in ways that conform to principles of the sort found in formal linguistics (e.g., Culler, 1976) are not the dominant perspective of the field. In more applied departments of some universities, such as those of animal husbandry, knowledge is also clearly organized around goals. Thus, knowledge clusters based on abstract characterizations of causal regularities are not the only way to understand knowledge, even in academic circles; but such clusters are a critical part of understanding much of the natural and social sciences.

The goal here is to explore how notions of discipline-based clusters emerge in development with a focus on those clusters most normally found in the natural and social sciences. Prior work has shown that even preschoolers do not think of adults as equally omniscient and sense some partitioning of knowledge (Lutz & Keil, 2003). Moreover, when discipline-based ways of clustering are pitted against object-based and goal-based ways of clustering, children under 10 years of age tend to prefer the goal- and object-based clusters over discipline-based ones (Danovitch & Keil, 2004). This preference raises the question of whether young children might not still have some sense of discipline-based clusters but have not yet granted them a privileged status.

The studies described here focus on disciplinary knowledge, arguably the most important way of understanding the division of cognitive labor and the way that may undergo the strongest developmental change during the elementary school years. Moreover, an awareness of discipline-based forms of knowledge may be most closely related to an emerging set of essentialist assumptions about natural kinds (Gelman, 2003). Thus, adult notions of essences tend to focus on the idea that there are deeper causal properties of natural kinds that determine kindhood and that correspond to legitimate areas of expertise (Gelman, 2003; Medin & Ortony, 1989; Putnam, 1975). This article asks whether much younger children do have some sense of discipline-based clusters when there are no conflicts with other ways of clustering and how that sense of discipline-based clusters develops and forms a basis for more sophisticated reasoning about kinds. We will suggest that young children's abilities to partially grasp some general causal and relational patterns do give them an early, albeit fragile, sense of discipline-based knowledge.

It might seem that children throughout the elementary school years, and perhaps even most adults, should be poor at discipline-based clustering. The vast majority of adults have had little or no exposure to disciplines such as chemistry, physics, biology, and psychology, and if exposure to those disciplines is critical to understanding the sort of phenomena a chemist might understand, then we would expect a layperson's intuitions about such matters to be minimal. In addition, the carving up of the disciplines into departments that populate the modern university is a recent historical event. One hundred and fifty years ago, the departmental organization of the university was very different from what it is today. Most of the natural sciences were absorbed into areas of philosophy and the social sciences were largely nonexistent (Pelikan, 1992). Moreover, although many adults may have at least heard of chemistry or biology in the press or in a high school class, elementary school curriculum does not usually break up science into such disciplines (Krajcik, Czerniak, & Berger, 1999; National Research Council, 2007). In addition, there is a tradition arguing that young children reason concretely in ways that could not be sensitive to the more abstract relations that form the basis of scientific disciplines (Bruner, 1960; Werner & Kaplan, 1963).

It is also the case that, in earlier periods of history, the notion of a polymath seemed more feasible. Many great figures of antiquity seemed to be experts on what would be regarded today as wildly diverse areas of knowledge. Aristotle, for example, seemed to be an expert on biology, physics, meteorology, memory, and perception, among other areas. By contrast, today, true polymaths seem highly unlikely given the ever-expanding depth and complexity of most areas of expertise. It might therefore seem that the ability to detect different areas of expertise is related to the relatively recent trend of only one area of expertise per individual. We doubt, however, that earlier generations had no awareness of such domains of understanding. After all, Aristotle did cluster his writings into books that often corresponded rather nicely with traditional disciplines (each of Aristotle's areas of expertise listed earlier were in distinct volumes). Similarly, libraries in the ancient world often had collections organized in ways that corresponded to many areas of expertise that might be recognizable today (Casson, 2001). Knowing one phenomenon that a person deeply understands may not reveal all of his areas of expertise, especially in more classical times, but it can reveal the coherent area of expertise governed by the principles underlying that phenomenon.

Another issue concerns findings of huge gaps in understanding in various domains. Ask people the details of how things work or why various natural phenomena exist as they do and one usually receives highly skeletal, fragmentary, or even contradictory bits of information (Chinn & Brewer, 1993; diSessa, Gillespie, & Esterly, 2004; Lawson, 2006). Similarly, people think they know far more about how many devices work because they falsely equate information that they can figure out in real time with information that they have actually represented (Keil, 2003; Rozenblit & Keil, 2002).

Failures in understanding, however, are usually revealed in attempts to offer explicit accounts of the details of mechanisms. If fully worked-out mental models and mechanisms are required to cluster knowledge by disciplines, then such a clustering would be in the rarified realm of a few brilliant “Renaissance” men and women. A closer look, however, suggests that far coarser, more skeletal senses of causal and relational patterns may still allow us to have an idea of discipline-based clusters.

Consider, for example, the following statement and question:

Adam knows all about why salt dissolves in water. Which of the following is Adam more likely to know about?

- a. Why dogs bark
- b. Why paper is flammable

Most adults answer that Adam is more likely to know b. They make this judgment with considerable confidence even though they themselves may know very little of the details of why salt dissolves, dogs bark, or paper is flammable. They believe that someone who knows all about the solubility of salt in water would have grasped the basic principles that inform combustibility more than barking. This inference may be mediated by a crude sense of regularities in the world governing the bonds that make up matter, regularities that are in turn connected to changes in states of matter.

Thus, if we look beyond explicit mechanistic knowledge to a more implicit sense of how things are patterned, a great deal more competence is suggested not only in adults but also in children. Even as researchers document glaring holes in people’s understanding of natural phenomena and complex artifacts, others document preschool intuitions of considerable subtlety concerning different sorts of patterns associated with domains such as living kinds (Hatano & Inagaki, 1999), physical mechanics (Spelke, Breinlinger, Macomber, & Jacobson, 1992), substances (Au, 1994), and folk psychology (Wellman, 1992). In many cases, a sense of abstract patterns in a domain seems to precede detailed mechanisms (Simons & Keil, 1995). Moreover, such patterns often seem to be in the form of schemas that represent expectations about how the world is causally structured, schemas that can override mere patterns of prior association (Keil, 2006; Mandler, 2004). Intuitive theories in this sense of coarse impressions of causal patterns seem to emerge very early and appear to be present in a variety of domains (Wellman & Gelman, 1998). Indeed, the spread of knowledge through a culture has been argued to be critically dependent on the presence of these early emerging domain-specific intuitive core or framework theories that ensure a necessary degree of stability and coherence to those domains (Sperber & Hirschfeld, 2004).

Young children's sense of these broad domains of regularities can also be seen in their judgments of "predicability" in which they see only certain classes of properties as being sensibly applicable to broad domains such as living things, artifacts, and psychological things (Keil, 1979, 1983). Their judgments of category mistakes, such as applying biological predicates to artifacts and solid object predicates to substances, suggest an early emerging implicit sense of broad domains of regularities that are delineated by patterns of natural language predicability. Those domains correspond roughly to disciplines such as biology, physical mechanics, and psychology. It therefore seems quite plausible that core domains of knowledge roughly analogous to biology, physics, and psychology might be sensed by young children at an abstract level that enables them to make inferences about the division of cognitive labor. In addition, although not explored nearly as much to date, other knowledge domains of the natural and social sciences might also be sensed and are also explored in the studies reported on in this article. Because patterns of predicability are similar across cultures, and because very young children and infants seem to have some sense of broad domains of causal and relational regularities, we would also expect that a sense of discipline-based areas of expertise should be present in a wide range of cultures and traditional nonindustrialized societies.

The presence of intuitive framework theories in young children, however, does not in itself guarantee that those who hold such theories can use them to infer the division of cognitive labor. Such inferences require additional assumptions that certain kinds of clusters of physical regularities predict clusters of knowledge and that it is plausible for a certain range of phenomena to be understood by one mind. These assumptions in turn may require a sense that one who grasps a few deep principles in domain should be able to use them generatively to understand all phenomena that are governed by those principles. Finally, there must be an appreciation of how discipline-based ways of clustering knowledge have privileged roles in reasoning about some types of problems, such as those about the essential natures of various kinds. For all these reasons, it is important to ask how children come to grasp discipline-based divisions of cognitive labor.

Given the example used earlier with solubility and combustion, adults and children might be able to use their implicit coarse senses of causal and relational structure to infer discipline-based clusters of knowledge long before they explicitly encounter those disciplines as such and long before they master the mechanistic details in those disciplines. A sense of these expertise domains may provide a critical basis for making inferences about the presence of essences in these domains even as children are unaware of their specific nature (Gelman, 2003). Assumptions about essences and deference to relevant experts on essences require a discipline-based sense of the division of labor as opposed to the other senses of expertise described above. Without a sense of disciplinary expertise, children would have no way to ground the large gaps in their knowledge.

As discussed earlier, quite young children, and even infants, do have some sense of the causal regularities that are associated with broad domains such as folk psychology, folk physics, and folk biology. We predict that, by the early elementary school years, children will sense enough about domain-specific causal patterns to be able to infer knowledge clusters roughly akin to academic disciplines. These inferences would require that children not only have a sense of causal patterns but also that they realize that deep expertise implies access to those causal patterns and thereby predicts that someone who fully grasps those patterns is also likely to

understand other phenomena that arise from those patterns, even if they are superficially quite different. However, because younger children will have much coarser and more skeletal insights into causal patterns, they should not be as aware of the relative richness of discipline-based knowledge and of its greater inductive fertility compared to other ways of classifying. As adults, we can see how someone who fully understands the principles of chemistry can understand a vast array of phenomena related to materials and their transformations. A young child may sense something about invisible microstructure being important to understanding substances and mixtures but may not sense enough of that structure to realize the special status of discipline-based knowledge. In addition, children would not be expected to directly sense a discipline that is historically recent (e.g., evolutionary theory), but they might well pick out phenomena roughly associated with that discipline by using a much older interpretative schema (e.g., adaptation). Similarly, children might pick up on a pattern not normally used to sort major disciplines but that does capture different kinds of deep causal regularities. Thus, there is not a perfect one-to-one mapping between perceived patterns of regularities and modern academic disciplines; but there may be enough correspondence to enable even young children to have some sense of who knows what. Studies 1 and 2 explore a broad sweep of disciplines to look for such patterns.

We would also expect the more fragmented causal and relational knowledge of young children to have more potential to lead children into making knowledge clusters that sometimes do not accord with patterns shown by older children and adults. For example, if a child sees a domain of “physics” as corresponding to causal patterns governing bounded bodies in motion (e.g., Michotte-type events such as launching, pulling and exploding; Scholl & Tremoulet, 2000; White, 2006), that child might mistakenly rule out static mechanics phenomena and include other phenomena in which bodies in motion are salient but incidental to the core domain of interest. We therefore predict that it should be possible, by using appropriate distractor “phenomena,” to uncover these simpler causal schemas and their distorting effects. Study 3 explores this issue.

If even young children are using their senses of causal and relational patterns to make inferences about knowledge clusters, then we predict that they should continue to do so when more surface cues to knowledge clusters are not relevant. This prediction is explored in Studies 4 and 5. Study 4 asks whether surface appearance of pictorial stimuli can explain performance. Study 5 asks whether sophisticated higher order associative relations between the words used in the stimuli descriptions can be used to predict clusters of expertise.

Finally, we expect that an increasing appreciation of the special status of discipline-based clusters will be related to changes in other cognitive abilities that might shed light on mechanisms of change. One such mechanism might be through cues from category labels. As one moves up the hierarchy of labels (e.g., dachshund > dog > animal), it becomes decreasingly plausible for there to be experts based on topic, access, or goals. Discipline-based forms of expertise, however, remain plausible at higher levels. Study 6 therefore explores the prediction that young children prefer discipline-based forms of categorization when higher level categories are used and that such a preference increases with age. A second mechanism may involve an emerging ability to realize that the same set of entities might be sorted into very different domains of expertise as a function of the kinds of causal relations that are invoked. Study 7 therefore examines the prediction that an emerging preference for discipline-based

clusters is related to a developing ability to reclassify expertise when new property types are encountered that suggest different kinds of causal and relational patterns as being central.

In short, we describe here a series of seven studies that explore the roots of a discipline-based form of understanding and how it becomes more robust with age. Above and beyond the specific predictions for each of the studies, the following questions guide our exploration:

- Are young children able to systematically access their intuitive senses of real-world causal regularities in a manner that enables them to infer clusters of knowledge in other minds? That is, can they tap into their intuitive theories of real-world patterns, sparse as they are, and use them to infer their corresponding domains of knowledge?
- How does the child's sense of discipline-based knowledge emerge during the elementary school years and become privileged relative to other ways of clustering knowledge? If some disciplines are distinguished earlier than others, how might such differences be explained?
- To what extent do children seem to make judgments about the division of cognitive labor by appeals to coarse causal schemas that comprise their intuitive theories as opposed to more associative grounds?
- If a disciplinary sense becomes stronger with increasing age, what are some possible mechanisms that might help foster a greater appreciation of the special nature of discipline-based forms of knowledge?

More generally, these studies ask whether discipline-based ways of construing expertise are relatively exotic ways of understanding knowledge that have arisen recently in history and emerge late in development or whether they are more foundational ways of understanding expertise that have much earlier roots and form a critical basis for children's emerging understanding about the causal complexity that underlies various natural kinds. We suspect the roots of discipline-based knowledge are very old in history and will appear very early in development.

## **2. Study 1: Developing notions of knowledge clusters**

Some divisions of academic labor are widespread in the modern university. Fields such as the social sciences and the natural sciences are labeled as such at most institutions even as there are disagreements about boundary disciplines like psychology. The natural sciences are usually divided up into physics, chemistry, and biology, and the social sciences into psychology, sociology, political science, and economics. Further common subdisciplines include molecular and evolutionary biology and cognitive and social psychology and are in many cases separate academic departments as well.

We do not argue that these are the only ways in which departments are organized. Many large land grant universities in North America have departments organized around goals rather than disciplinary principles, such as departments of hotel management or poultry production. Moreover, some groups seem to cluster less because of a common set of principles and more because of a common object of study, with at least some cognitive science groups being organized on that basis. Disciplinary principles are simply one way of clustering that

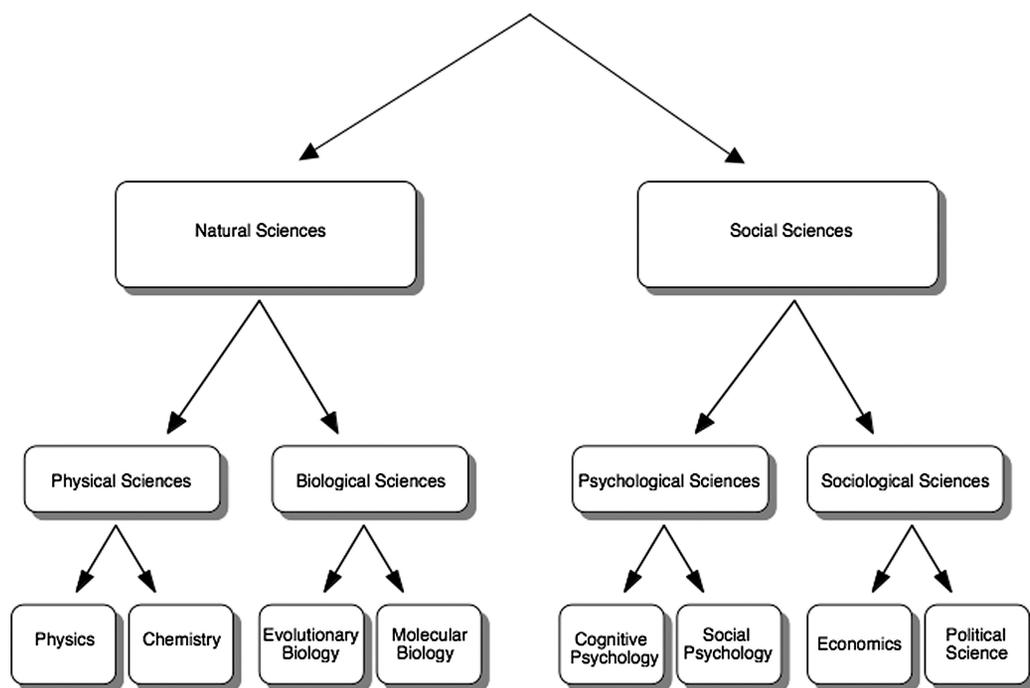


Fig. 1. A hierarchical organization of academic disciplines.

is especially common in the “basic” sciences and some social sciences and that corresponds most closely to natural kinds.

As the initial step in exploring knowledge about disciplines we decided to use the hierarchical structure shown in Fig. 1. The largest cut occurs between the natural and the social sciences. The next level of physical and life sciences is commonplace. The division between psychological and sociological sciences, while not often labeled as such, does capture a frequently made distinction in academic circles. Finally, the bottom level corresponds to either common academic departments, such as physics and chemistry, or common subdivisions, such as cognitive and social psychology.

This hierarchical structure allows exploration of several levels of analysis as well as providing for a balanced symmetrical stimulus set that allows for a cleaner experimental design. It is not meant in any way to represent the “correct” map of the academic disciplines; rather, it represents one map that is assumed to correspond to clusters of lawful regularities in the real world. Disciplines and subdivisions that are closer together in the lowest nodes of the hierarchy are those that might be assumed to have the largest overlap of causal regularities and principles or, more pragmatically, to have more jointly appointed faculty. We are, of course, not saying that young children have any sort of explicit awareness of some of these contrasts, especially the lower ones such as those between evolutionary and molecular biology and between cognitive and social psychology, both of which are relatively recent divisions of adult knowledge. But it may well be that the different patterns of causation and relational structure that distinguish these subdivisions can be sensed by young children and used to make inferences about domains of knowledge. For example, a child might sense that there

are patterns in biology that seem to rely on teleological/functional contrasts (evolutionary biology) and others that rely on essences (molecular biology).

With this tree as a basis for experimental design, in the experiments that follow it was necessary to develop a set of phenomena that were typical of each discipline and which, at the same time, would both capture deep principles and be clearly sensible to a young child. The methods for constructing the materials are described in the Stimuli section. Once stimuli suitable for eventual use with children were developed, we presented the knowledge-clustering task to adults.

## 2.1. Method

### 2.1.1. Participants

Fourteen male and 19 female university undergraduates participated in this study along with 165 children: 53 kindergartners, 61 second graders, and 51 fourth graders (mean ages, 6 years 2 months; 8 years 2 months; 10 years 2 months). The design called for 48 children in each age group and 32 adults, but a desire to include all interested children in a classroom and a decision to accommodate adults who signed up beyond the list limit led to slightly larger numbers of participants. Adults were run individually with each session lasting approximately 7–12 minutes. Children were interviewed individually outside of their classroom with each session lasting approximately 20 minutes. There were roughly equal numbers of males and females in each condition. All were from the greater New Haven area and reflected the general demographics of that population (approximately 75% White, 13% African American, 6% Asian, 6% other, with most children being of middle-class backgrounds). None of the children who participated were from families in which their parents were university-level faculty or researchers in one of the traditional academic disciplines.

### 2.1.2. Materials and procedure

*2.1.2.1. Stimuli:* Four Yale University undergraduates, naïve as to the purpose of the experiment and the predictions, were asked to generate a set of phenomena central to the disciplines of physics, chemistry, molecular biology, evolutionary biology, cognitive psychology, social psychology, political science, and economics. They were told to select phenomena for which the deep understandings would require a grasp of the central principles of a discipline. At the same time they were told that the phenomena themselves should be described in very simple terms that even a young child could understand. The children should be able to understand what the phenomenon was even as they and most adults might be completely unable to explain why the phenomenon occurred as it did. A few examples were offered.

Five to eight sample phenomena were generated for each of the disciplines. The experimenters then edited examples to simplify lexical items, to minimize overlap of common terms, to standardize syntax as much as possible, and to make the lengths as equivalent as possible. The experimenters picked what seemed to be the four simplest and clearest examples for each discipline. The actual examples for evolutionary biology were all cases of obvious adaptations, only a subset of the field. Similarly, the examples of molecular biology were all cases of visible effects of biochemical and genetic processes. Physics as well was restricted to cases of mechanics.

Table 1  
Disciplines used to construct triads for both near and far comparisons

	Physics	Molecular Biology	Social Psychology	Economics
Near comparison	Chemistry	Evolutionary biology	Cognitive psychology	Political science
Far comparison	Cognitive psychology	Political science	Chemistry	Evolutionary biology

Using these edited sets of phenomena, a set of triads was constructed. Each triad was of the form “This expert knows all about X. Would they know more about Y or Z?” One member of the Y, Z pair was from the same discipline as the first fact (X), or was discipline consistent, and one was from a different discipline, or was discipline inconsistent. For example: if X was “This expert knows all about why people make a shopping list when they are going shopping later” (cognitive psychology), the discipline consistent choice was (Y) “Why you can’t remember things from when you were a baby” (cognitive psychology), and the discipline inconsistent choice was (Z) “Why a big, heavy boat takes a really long time to stop” (physics).

For every discipline there was a “near” comparison with a more closely related discipline in the hierarchy and a “far” comparison for more distantly related disciplines in the hierarchy. Table 1 illustrates the comparisons. The disciplines across the top of the table indicate which phenomena was the basis for the comparison as well as the discipline-consistent choice.

Half of the triads had discipline-inconsistent items from near disciplines (e.g., chemistry when X was from physics) and the other half were constructed using phenomena from far disciplines (e.g., cognitive psychology when X was from physics). A total of 32 triads were created, four for each of the eight item types: 16 near triads (a) molecular biology/evolutionary biology, (b) economics/political science, (c) physics/chemistry, (d) social psychology/cognitive psychology, and 16 far triads: (e) physics/cognitive psychology, (f) social psychology/chemistry, (g) molecular biology/political science, (h) economics/evolutionary biology. The triads were split into two conditions, A and B; each one contained two near item types and two far item types for a total of 16 triads in each condition.

Condition A consisted of four molecular biology/evolutionary biology triads, four economics/political science triads, four physics/cognitive psychology triads, and four social psychology/chemistry triads, and condition B consisted of four physics/chemistry triads, four social psychology/cognitive psychology triads, four molecular biology/political science triads, and four economics/evolutionary biology triads. Each participant was randomly assigned to one of the two conditions. The full set of triads is shown in Appendix B.

**2.1.2.2. Procedure: Adults**—Each participant was given 16 groups of sentences (triads) in a random order. Participants were told that each group of sentences involved a different expert who knows about a certain fact. They were then asked to choose which of the two other facts the expert would be more likely to know about. The notion of experts was used to emphasize that full mastery of knowledge in the area was involved.

**Children**—At the beginning of each interview, the experimenter discussed with the child the definition of an expert. Children were asked first if they knew what an expert was. If they said “yes,” they then were asked to give an example of an expert. The experimenter continued with two other examples of experts, a doctor and a car mechanic, explaining that a doctor is an expert because he knows all about how to fix people and make them feel better and that a

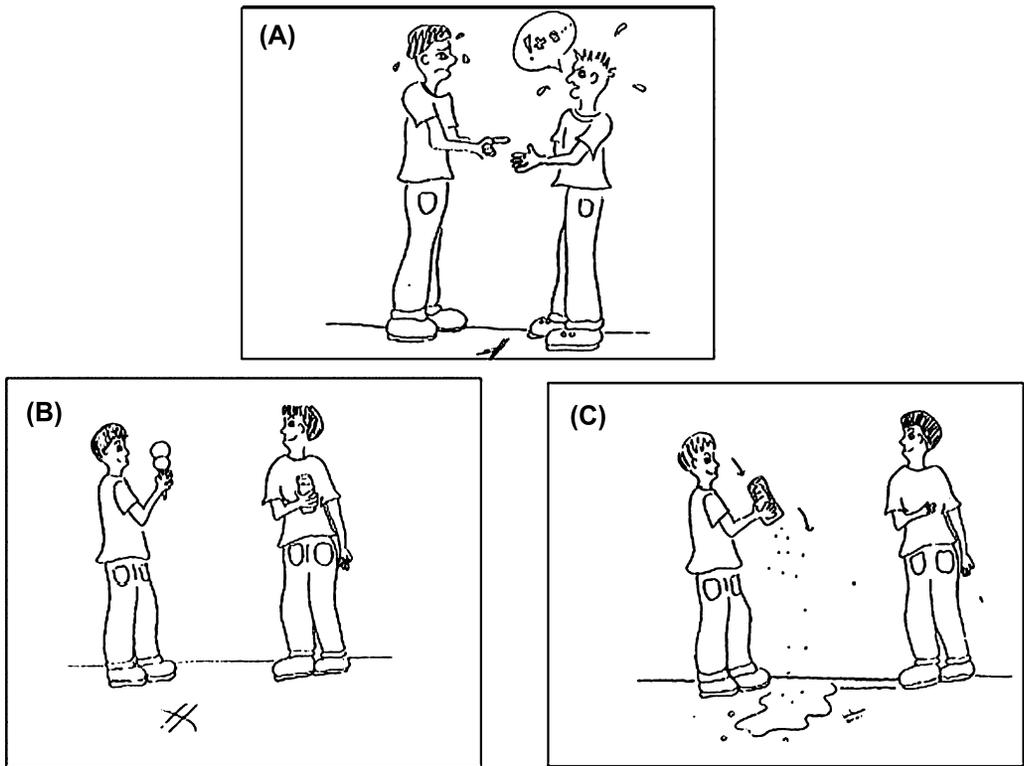


Fig. 2. Example of pictures used in triads.

car mechanic is an expert because he knows all about how to fix cars and make them work again (Lutz & Keil, 2003). If the child said that he did not know what an expert was, then the experimenter explained that “an expert is someone who knows a whole lot about something, but they don’t know everything” and then continued with the examples above. Each child was given two training items. The first training item was used to ensure that children understood the format of the task. If they answered incorrectly the experimenter gently corrected them and they were given an explanation as to why the other choice was a more appropriate answer. The second training item was more critical: it required that the children actually think about the statements and infer what the expert knows about (see Appendix A for full script). If they failed the second item, their data was excluded since they did not understand the concept of an expert. (Two children were excluded, both from the kindergarten age group). During the task, children were not corrected for discipline-inconsistent answers but were asked for each item why they believed that the expert would know more about the phenomenon that the child had selected.

Each statement was accompanied by a drawing that depicted what was happening (for an example, see Fig. 2). Drawings were matched on content (e.g., if there were two people in the first drawing, there were also two people in the other two drawings). Children were told that the pictures were there to help them remember what the expert knew about and that there were

no clues in the drawings. A study with adults and children (Study 4) was conducted to ensure that the participants were not answering based on mere perceptual similarity of the drawings.

By way of illustration, the triads accompanying the pictures used in Fig. 2 are as follows:

Adam knows all about why:

People sometimes fight more when they are tired (A)

What else is he more likely to know?

Why people smile at their friends when they see them (B)

Or

Why salt on people’s icy driveways makes the ice melt sooner (C)

2.2. Results

If participants chose the discipline-consistent answer their responses were coded “1” and if they chose the discipline-inconsistent answer their responses were coded “0.” A total score, ranging from 0 to 8, was tallied for each participant for both the near triads and for the far triads. Fig. 3 shows the mean percentage of discipline-consistent answers for near and far comparisons for all age groups.

A 2 × 4 repeated measures ANOVA was used to analyze the number of discipline-consistent clusters, with distance (near vs. far) as the within-subject factor and age (K, 2, 4, Adults) as the between-subject factor. Effect size estimates were computed using partial  $\eta^2$ . The results

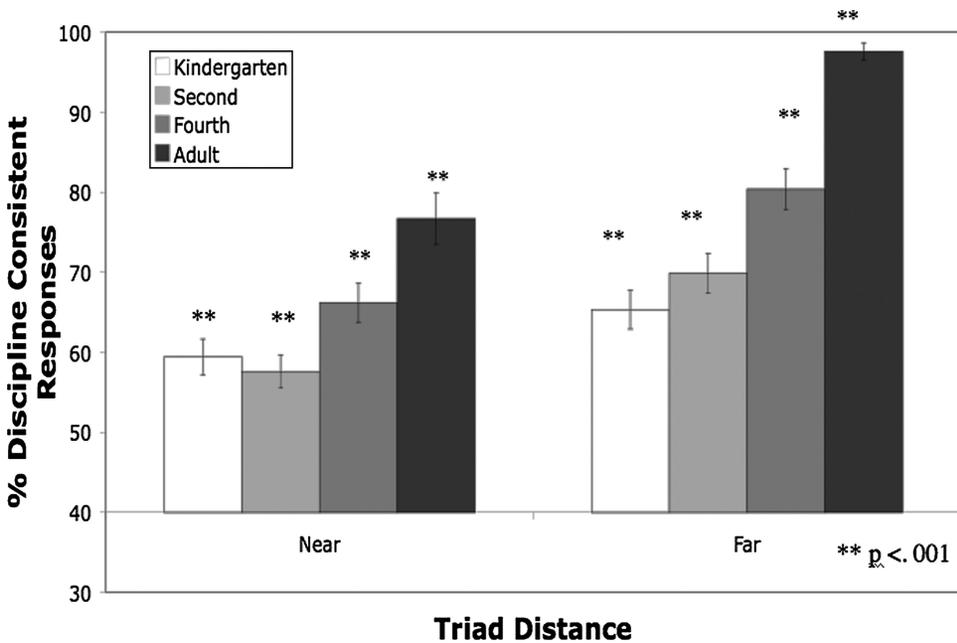


Fig. 3. Mean percentage of discipline-consistent answers given for all near items and all far items by age. Asterisks indicate clusterings by discipline exceeding chance levels. Bonferroni corrections for multiple comparisons are used for all similar analyses throughout the article. Significance levels represent probabilities at or below the corrected values.

Table 2  
Mean (standard deviation) number of discipline consistent answers for each of the triad types by age (range = 0–4)

	Age			
	Kindergarten	Second	Fourth	Adult
Near				
MbioEbio	2.86 (.95)*	3.06 (.83)*	3.40 (.82)*	2.60 (1.1)
EconPolisci	1.93 (.92)	1.24 (1.2)*	1.64 (1.1)	2.67 (1.2)
SocCog	1.96 (.96)	2.18 (.91)	2.77 (.91)*	3.60 (.63)*
PhysChem	2.75 (1.0)*	2.79 (.74)*	2.77 (.82)*	3.33 (.90)*
Far				
MbioPolisci	2.67 (.92)*	2.89 (.92)*	3.31 (.84)*	3.87 (.35)*
EconEbio	2.83 (1.2)*	2.89 (1.1)*	3.12 (1.1)*	3.87 (.35)*
SocChem	2.34 (1.1)	2.61 (.93)*	3.24 (.93)*	3.93 (.26)*
PhysCog	2.66 (.94)*	2.82 (.85)*	3.20 (1.0)*	3.93 (.26)*

Note. Probabilities < .002 are used to indicate significance, assuming a conservative Bonferonni correction for multiple comparisons.

\*  $p < .002$ .

indicate a main effect of distance,  $F(1,192) = 70.61$ ,  $p < .001$ ,  $\eta^2 = .269$ , suggesting that participants made more discipline consistent clusters in the far condition ( $M = 6.06$ ) than in the near condition ( $M = 5.07$ ). There was also a main effect of grade,  $F(3,192) = 29.68$ ,  $p < .001$ ,  $\eta^2 = .317$ , showing more discipline-based clusters in older children and a significant distance by grade interaction,  $F(3,192) = 3.37$ ,  $p < .02$ ,  $\eta^2 = .050$ , suggesting a greater contrast between near and far disciplines in older children (see Table 2 for means).

Although the kindergarteners and second graders performed worse than the fourth graders, a one-sample  $t$ -test (using a Bonferroni correction) over all contrast types shows that both the kindergarteners,  $t(52) = 6.34$ ,  $p < .001$ , and second graders,  $t(60) = 8.05$ ,  $p < .001$ , performed well above chance levels (mean = 4.0) on the far comparisons as well as on the near comparisons, kindergarten,  $t(52) = 4.28$ ,  $p < .001$ ; second grade,  $t(60) = 3.70$ ,  $p < .001$ .

As can be seen in Table 2, separate one-sample  $t$ -tests, using a Bonferroni correction for multiple comparisons, showed that each of the triads in the far condition were above chance levels with the exception of the social psychology/chemistry triads at the kindergarten age group. The near comparisons, however, had fewer triads reaching levels above chance, with the second graders actually choosing the discipline inconsistent answer significantly more in the economics/political science triads.

### 2.3. Discussion

For the adults, two patterns are clear from Study 1. First, adults agree strongly on how to cluster expert understandings of natural phenomena even when they themselves have very little grasp of those expert understandings. Second, the difference between near and far disciplines,

as represented in the tree, has psychological reality as revealed by the greater difficulty in achieving consistency on the triad task for “near disciplines.”

The high levels of performance on this task, even with minimal deep knowledge, raises the central question of how adults do so well. This issue is further developed in the studies that follow, but comments from some adult participants suggest that they are sometimes invoking simple causal schemas that are shared by two of the three phenomena.

Knows all about usefulness and how that influences something’s value. (economics)

The first one knows about the transfer of energy from motion to force and the results. Harder jumping makes you go harder, swinging the hammer makes nailing harder. (physics)

They are both issues of increasing energy input to increase output. (physics)

Often, the explanations offered describe a global relationship that is central in a domain even as they don’t describe the details in ways that are commonly used by experts (as in the physics case). This way of discussing the relations suggests that they might be accessible to very young children as well.

Several patterns emerge from the child portion of this study as well. First, even the youngest children consistently perform at above chance levels for three of the four “far” contrasts, suggesting that they have some intuitions as to how knowledge might be distributed in the minds of others. Their responses indicate clear biases about how pieces of knowledge might be clustered.

Second, there is some indication in the children’s responses that they are aware of underlying relations that organize a domain, relations to which the expert has privileged access. Consider the following responses from some fourth grade participants:

They both have to do with feelings. (social psychology)

The ball comes down because of gravity and the cart is light because of gravity. (physics)

The hot dog one because it involves selling. (economics)

Although the majority of children were not able to say much to justify their responses, when they did, they would often refer to relations or higher order abstract concepts that seem to capture disciplinary insights. The child who clustered “people feel bad if their best friend doesn’t play with them” and “why grown-ups fall in love” and said it was because “They both have to do with feelings” was noting that emotional states are central to expertise in the domain. The child who referred to gravity transcended the vertical and horizontal orientations to see a deeper principle. Finally, the child who clustered economics questions on the basis of their both involving “selling” did so despite no specific mention of selling in the triad items. Because the comments were fairly infrequent, it was not feasible to conduct a quantitative analysis of comment types.

Since intuitions about disciplinary knowledge clusters often arise from tacit knowledge, even in adults, difficulty in articulating reasons for the clustering should not be surprising. Study 3 explores this implicit knowledge in a different manner not requiring explicit justification.

Third, at all ages there was some psychological reality to the near/far contrast shown in Table 1, with a greater ability to cluster the two same discipline pairs when pitted against a far contrast. This near vs. far difference, however, was largely carried by the cognitive

vs. social psychology contrast and the economics vs. political science contrast as even the youngest children seemed to distinguish chemistry from physics and evolutionary biology from molecular biology. It is important to remember that the evolutionary versus molecular biology contrast refers to the typical adult disciplinary labels. Given the stimuli and the children's responses, the actual contrast they used seem to be between adaptive relations in biology and physiological ones.

Finally, there is the peculiar result that younger children actually did worse than chance on the near political science/economics contrast. A closer inspection of the stimulus items suggests that the children were picking up on a different contrast that was unwittingly part of the stimuli, namely the difference between macro and micro processes and institutions. Some political science and economics items involved large groups or institutions such as cities or companies, whereas others involved individuals and their behaviors. By chance, this macro vs. micro contrast formed clusters that paired up several political science questions with economics questions.

The occasional mention by children of notions such as "feelings" or "selling," although suggesting use of some sort of abstract relation does not in itself make clear the nature of those relations. They might be understood primarily in causal terms, such as that there are certain patterns of how emotions lead to actions; or they might be understood as merely relations that are unique to a domain, such as that feelings are somehow common to both phenomena in one domain and not to another. Even more superficially, children might be basing their judgments on noticing which words are likely to occur near each other in discourse; i.e., if they have heard the word "feelings" used in discourse about certain sets of phenomena, then they will be considered in common areas of expertise. These alternatives are explored in a series of studies presented later in this article.

### **3. Study 2: Reverse triad**

Study 1 showed that children can successfully group two items from the same discipline in a triad where one item is from a different discipline than the other two. The ability to discern such clusters raises the question of how the children succeed. Consider two possibilities. When children, for example, group together two physics items and exclude a chemistry item, their grouping might arise from a rudimentary understanding of both of the domains of chemistry and physics and how they contrast. Alternatively, their grouping might result from a good skeletal theory of the physics domain, with little understanding of chemistry. Children might simply match the two physics items together because they fit a familiar schema and see the chemistry item as not physics.

To explore these alternatives in more detail, and to replicate the results of Study 1, we constructed a new set of stimuli. The new reverse triad stimuli reversed the direction of the discipline-consistent and the discipline-inconsistent contrasts. For example, in a Study 1 item a child would hear about a physics expert and would then have to choose whether the physics expert would know more about another physics item or a chemistry item. In the reverse triad version of the stimulus, a child would hear about a chemistry expert and then would have to choose whether the chemistry expert would know more about another chemistry item or

a physics item. If children were much more comfortable with their grasp of one domain in a two-domain contrast, it seemed likely that they would do better in cases where two of the three items in the triad came from that domain. Thus, if children in Study 1 easily saw the physics/chemistry contrast with two physics items and one chemistry item in the triad but did considerably worse in Study 2 with one physics item and two chemistry items, such a pattern would suggest that a better grasp of physics was responsible for success on the task.

By examining the changes in performance between Study 1 and Study 2 for all near-far shifts, it should be possible to both evaluate the replicability of the results and to get some hints at the basis for children's judgments.

### 3.1. Method

#### 3.1.1. Participants

Seventeen college students participated in this study, along with 93 children: 32 kindergartners, 32 second graders, and 29 fourth graders (mean ages, 5 years 5 months; 7 years 6 months; 9 years 4 months. The design called for 32 children in each age group and 18 adults; practical recruitment issues led to a slight departure from these numbers.) Adults were run individually with each session lasting approximately 7–12 minutes. Children were interviewed individually outside of their classroom with each session lasting approximately 20 minutes. There were roughly equal numbers of males and females in each condition. All were from the greater New Haven area and reflected the general demographics of that population.

#### 3.1.2. Materials and procedure

*3.1.2.1. Stimuli:* A new stimulus set was constructed for the study. The stimuli were similar to those used in Study 1, except that the discipline-consistent and the discipline-inconsistent items for the stimuli were reversed. For example, if in Study 1 the triad consisted of a physics fact followed by another physics fact and a chemistry fact, in Study 2 the triad consisted of a chemistry fact followed by a physics fact and another chemistry fact (see Appendix C for full list of triads).

Note that while the original triad physics/chemistry contrast required two physics items and one chemistry item, the reverse triad chemistry/physics contrast required two chemistry items and one physics item. Therefore, the change in the stimuli required creating additional knowledge items for the original expert discipline (e.g., the number of physics items was doubled). New phenomena and drawings were created to use in addition to the existing ones for chemistry, cognitive psychology, political science, and evolutionary biology. These new items also served to provide a broader replication of Study 1 with new items.

*3.1.2.2. Procedure:* Except for the change in stimuli, the procedures for this study were otherwise identical to Study 1. The same introduction and training were used with the children and the same criteria for exclusion were also used. Again, children were not corrected for discipline-inconsistent answers but were asked, for each item, why they believed that the expert would know more about the phenomena that the child had selected. One kindergarten child was excluded for failure to pass the training items.

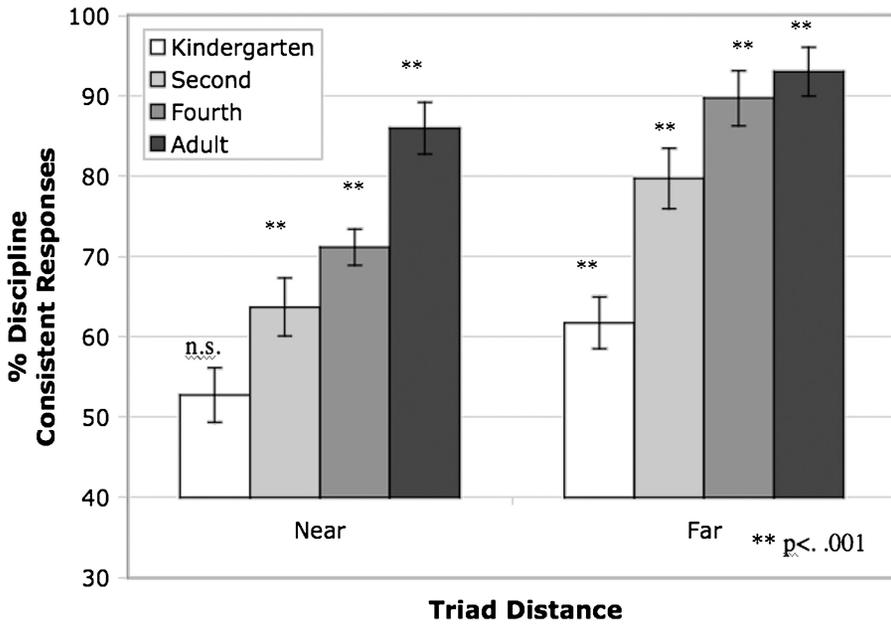


Fig. 4. Mean percentage of discipline consistent answers given for all near items and all far items by grade on the reverse triad task. Significance levels indicate mean scores were above chance.

### 3.2. Results

If participants chose the discipline-consistent answer their responses were coded “1” and if they chose the discipline-inconsistent answer their responses were coded “0.” A total score, ranging from 0 to 8, was tallied for each participant for both the near triads and for the far triads. Fig. 4 depicts the mean percentage of discipline-consistent answers for the combined near and far item types by age group.

A  $2 \times 4$  repeated measures ANOVA was used to analyze the number of discipline-consistent clusters, with distance (near vs. far) as the within-subject factor and grade (K, 2, 4, adult) as the between-subject factor. The results indicate a main effect of distance,  $F(1,105) = 24.59$ ,  $p < .001$ ,  $\eta^2 = .190$ , again indicating that participants are making more discipline-consistent clusters in the far condition ( $M = 6.32$ ) than in the near condition ( $M = 5.26$ ). There was also a main effect of grade,  $F(3,105) = 29.83$ ,  $p < .001$ ,  $\eta^2 = .460$ , indicating that the older children are making more discipline clusters than the younger children, but no interaction between distance and grade,  $F(3,105) = 1.15$ ,  $p = .331$ ,  $\eta^2 = .032$ .

Again, although the kindergarteners and second graders performed worse than the fourth graders, one-sample t-tests (using a Bonferroni correction) indicate that both the kindergarteners,  $t(31) = 3.64$ ,  $p < .001$ , and second graders,  $t(31) = 7.91$ ,  $p < .001$ , performed above chance levels (mean = 4.0) on the far comparisons taken as a whole. The second graders also performed above chance levels on the near comparisons,  $t(31) = 3.79$ ,  $p < .001$ ; however, the kindergarteners did not,  $t(31) = .80$ ,  $p = .43$ .

Table 3  
Mean (standard deviation) number of discipline consistent answers given for each of the triad types by age

	Age			
	Kindergarten	Second	Fourth	Adult
Near				
EbioMbio	1.79 (.98)	2.44 (1.2)	2.88 (.89)*	3.70 (.48)*
PolisciEcon	2.14 (1.1)	2.69 (.70)*	2.50 (.52)*	3.00 (.67)*
CogSoc	2.17 (1.1)	2.63 (.96)	3.38 (.65)*	3.83 (.41)*
ChemPhys	2.28 (.83)	2.56 (1.1)	2.69 (.48)*	3.33 (.82)
Far				
PolisciMbio	2.72 (.96)*	3.69 (.48)*	3.54 (1.1)*	3.67 (.82)*
EbioEcon	2.67 (1.0)	3.25 (.93)*	3.31 (1.0)*	3.67 (.52)*
ChemSoc	2.29 (.83)	2.56 (1.1)	3.75 (.45)*	3.60 (.70)*
CogPhys	2.07 (1.1)	3.25 (1.1)*	3.69 (.60)*	3.90 (.32)*

Note. Probabilities < .002 are used to indicate significance, assuming a conservative Bonferonni correction for multiple comparisons.

\* $p < .002$ .

As can be seen in Table 3, one-sample t-tests revealed that the kindergarteners were at chance levels on all but one of the triad types; however, the second and fourth graders were above chance performance on most of the far items. For the near items, only the fourth graders and adults were above chance, whereas the kindergarteners and second graders were at chance levels.

Using independent samples t-tests, Study 1 was compared with Study 2. The results suggest that children of all ages may have had insight into both domains involved in the contrasts. Thus, reversing the pairings of domains had no significant impact on judgments except in the case of molecular biology/evolutionary biology–evolutionary biology/molecular biology,  $t(131) = 3.98$ ,  $p < .001$ , and economics/political science–political science/economics,  $t(131) = -4.71$ ,  $p < .001$ . In these two instances, there was better performance when the molecular biology and the political science items were the two-item pair of the triad.

### 3.3. Discussion

Two key findings emerge from Study 2. First, it largely replicates the findings of Study 1. In the far condition, the second graders sorted most phenomena appropriately and even the kindergartners were capable of sorting at least one set of phenomena into categories corresponding to fundamental disciplines (political science and molecular biology). This was, however, a weaker performance than for the kindergartners in the first study, who sorted three out of four triads appropriately. The advantage of far over near contrasts was replicated. Finally, there was also once again a clear increase in the robustness of discipline based sorting with age.

The second key finding concerns the effects of reversing which domain had the two examples in each triad set. The reversals had relatively few effects on performance, suggesting that insights into the nature of both domains were contributing to performance on this task. The three exceptions may have involved the molecular biology/evolutionary biology, political science/economics, and physics/chemistry, where a somewhat greater insight early on into phenomena arising from principles of molecular biology, political science, and physical mechanics is suggested relative to evolutionary biology, economics, and chemistry.

#### 4. Study 3: Schema task

Children as young as 5 years of age are able to sometimes cluster knowledge in ways that seem to illustrate an appreciation of the kinds of basic disciplinary clusters of the natural and social sciences in the modern university. Most elementary school students have had little or no exposure to the traditional disciplines either by their names or by curricular content organized around such topics as chemistry, physics, and molecular biology. We have argued, however, that one way to accomplish such sortings of knowledge requires no detailed, explicitly taught knowledge of a discipline but rather use of more implicit interpretative schemas for different aspects of the world. Adults and children alike may believe that quite abstract patterns apply to a domain without knowing many of the details at all. They might further believe an expert in one of those domains would know all about the details linked to such schemas and therefore know all about the details of other phenomena governed by the same schema.

The use of schemas, however, can only approximate the richer causal and relational structure of the world. The schemas are a necessarily sparse encoding of complex relations (Keil, 2003). This sparseness suggests a method for testing the content of the schema: by creating stimuli that fit the schema but not the discipline (+schema/–discipline cases) and those that fit the discipline but not the schema (–schema/+discipline cases).

If a child groups (+schema/–discipline) cases with the (+schema/+discipline) cases, we can infer that the child is using that particular schema to cluster knowledge. If, by contrast, the child groups the (–schema/+discipline) cases with the (+schema/+discipline) cases, we can infer the child is using some other, perhaps deeper, understanding of the phenomena.

For example, suppose a child sees the domain of mechanics as involving bounded objects in motion and the ways in which they behave in cases of direct contact with other bounded objects. In fact, this might be quite close to the kinds of mechanical causation that are thought to be directly perceivable, such as launching, pulling, and bouncing (Scholl & Tremoulet, 2000; White, 2006). The child might assume that there are sets of laws, which they do not know, that govern those kinds of relations. Their schema captures the kinds of relations but not the underlying laws.

This, in fact, is a fairly good heuristic for capturing the domain of mechanics, but if used exclusively, some mistakes will be made. Some cases of mechanics involve no bounded objects in motion but, rather, static situations, such as the forces that explain how a suspension bridge stands up or that explain why friction keeps a block from sliding down an inclined plane. Other situations do involve bounded objects in motion but the critical relations to be explained

do not involve mechanics. Thus, knowing why a fired bullet cannot be seen is a problem in perception, not mechanics, even though it showcases a bounded object in motion.

In Study 3 we described an individual as knowing how or why something was the case where the phenomenon was both +schema and +discipline. We then presented +schema/−discipline and −schema/+discipline alternatives to choose between. The extent to which a child was influenced by a +schema/−discipline example was taken as support for the use of that schema.

To construct the schemas, all the comments in Studies 1 and 2 were examined for mention of causal schemas in justifications for clustering judgments. As most of the children in Studies 1 and 2 did not verbalize the reasons for their judgments, the positive schemas could only be rough approximations based on the few comments that were given. In addition, since different schemas were sometimes mentioned, only the most commonly mentioned and simple one was used in each condition. For those reasons it was predicted that children would not be completely swayed by any one schema in the +schema/−discipline case, but that the schema would cause a significant drop in clustering by discipline. It was also predicted that as their schemas became more differentiated and sophisticated, older children would be less and less swayed by the +schema/−discipline cases, which used sparse simple schemas of the sort thought to be more common in younger children.

#### 4.1. Method

##### 4.1.1. Participants

Eleven college students participated in this study along with 102 children: 26 kindergarteners, 26 second graders, 25 fourth graders, and 25 sixth graders (mean ages, 5 years 8 months; 7 years 7 months; 9 years 7 months; 12 years 1 month. The design called for 25 children in each age group and 16 adults, but a desire to include all interested children in a classroom and recruitment issues with adults led to a modest departure from these numbers.) Sixth graders were included in this study in the event that the misleading schema delayed the development of discipline-based clusters. Adults were run individually with each session lasting approximately 7–12 minutes. Children were interviewed individually outside of their classroom with each session lasting approximately 20 minutes. There were roughly equal numbers of males and females in each condition. All were from the greater New Haven area and reflected the general demographics of that population.

##### 4.1.2. Materials and Procedure

4.1.2.1. *Stimuli*: For this study, we focused solely on the far comparisons. These were the clearest contrasts at all ages and were therefore where the strongest possible effects of the schema manipulations might be observed. The children's original statements from Studies 1 and 2 were analyzed and coded for simple schemas. Based on the children's responses, candidate schema lists were made for each of the eight disciplines. There were distinct schemas that could be found in all of the statements for each of the disciplines (e.g., for physics, bounded objects in motion). Three adult judges conferred on what single schema seemed most dominate for each discipline. These schemas were then used to create new statements and new triads were formed. The expert statement (X) contained one of the original statements from Study 1; however, in this case rather than have the children just

Table 4

List of schemas used in Study 3 to examine the extent to which schemas could cause a drop of discipline-based responses for far comparisons

Discipline	+schema/ +discipline	–schema/ +discipline	+schema/ –discipline
Physics	Bounded objects in motion/physics	Bounded objects not in motion/physics	Bounded objects in motion/cognition
Molecular biology	Consequence is internal/molecular biology	Consequence is visible/molecular biology	Internal consequences/political science
Social psychology	Dyadic interactions/social psychology	Self-reinforcing and self-monitoring behaviors/social psychology	Dyadic interactions/chemistry
Evolutionary biology	Adapted visible function/evolutionary biology	Function is more inferred (evolutionary history)/evolutionary biology	Adapted visible function/economics

choose between two academic disciplines (Y and Z), schemas were pitted against disciplines. If the expert knows all about a physics fact (e.g., why a hammer drives a nail better when you swing the hammer faster), we gave the children two choices: (a) a cognitive psychology fact that had something to do with bounded objects in motion attached to it (e.g., why people do better catching a ball with two eyes open instead of one) and (b) a physics fact that had little to do with bounded objects in motion (e.g., why big bridges need really big supports). In this manner, after receiving the original expert statement, each child received two choices: a schema-consistent but discipline-inconsistent case and a schema-inconsistent but discipline-consistent case. A list of the schemas used is shown in Table 4. Sixteen triads were constructed (see Appendix D for full list of triads).

*4.1.2.2. Procedure:* The procedure for this study was again identical to Study 1. The same introduction and training were used with the children and the same criteria for exclusion were also used. Again, children were not corrected for discipline-inconsistent/schema-consistent answers but were asked why they believed that the expert would know more about the phenomena that the child had selected.

## 4.2. Results

If participants chose the discipline-consistent/schema-inconsistent answer their responses were coded “1” and if they chose the discipline-inconsistent/schema-consistent answer their responses were coded “0.” Fig. 5 depicts the mean percentages of discipline-consistent answers for each of the item types by age group. It also plots, for grades K, 2, and 4, the results of Study 1 for the purposes of comparison (there were no grade 6 children in Study 1).

The use of misleading schemas did seem to cause reductions in discipline based clusters at all ages. A  $2 \times 3 \times 4$  repeated measures ANOVA was used to determine the extent

of this reduction, with study (1 vs. 3) and grade (K, 2, 4) as the between-subject factors and domain pair (mbio/polisci, phys/cog, soc/chem, ebio/econ) as the within-subjects factor. There was an overall drop in the use of discipline based clusters from Study 1 ( $M = 11.49$ ) to Study 3 ( $M = 9.12$ ),  $F(1,149) = 70.64$ ,  $p < .001$ ,  $\eta^2 = .322$ . This drop was found in

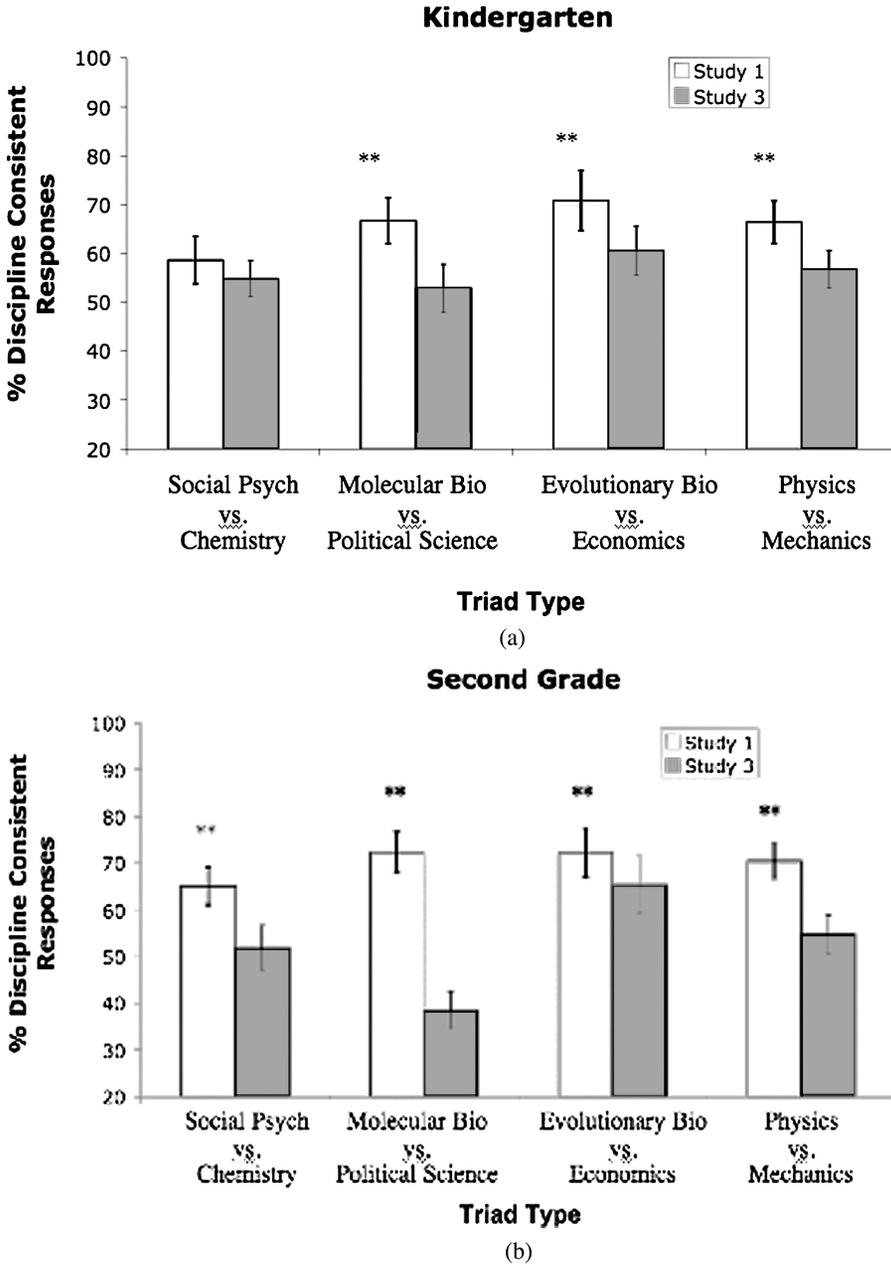


Fig. 5. Mean percentage of discipline-consistent choices by grade. The graphs for grades K, 2, and 4 include, for comparison purposes, means from Study 1 as well (there were no sixth-grade children in Study 1). \*\*Corresponds to all cases where means scores were above chance at a level of  $p = .003$  or less. (Continued on next page).

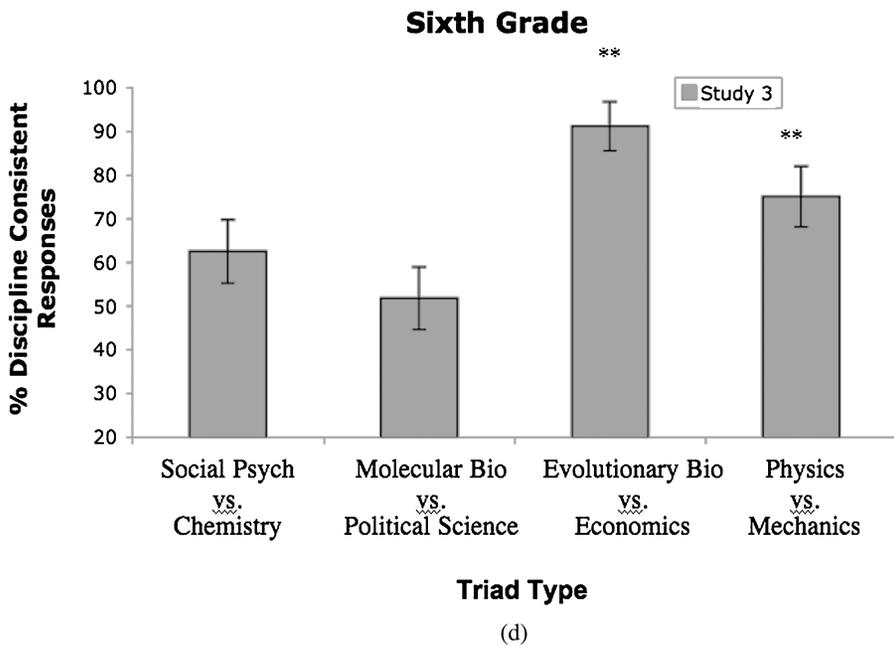
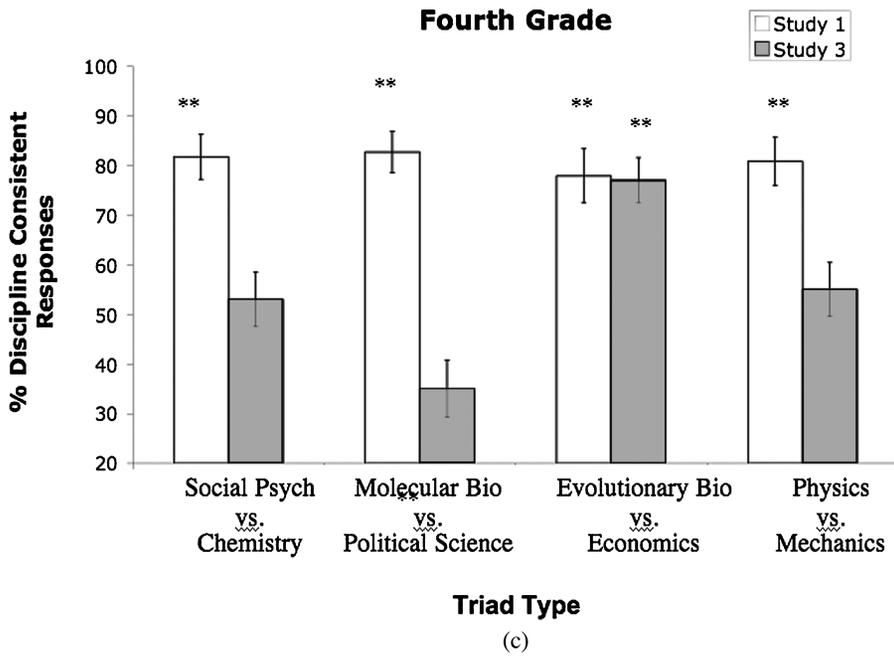


Fig. 5. (Continued).

three of the four comparisons (social psychology vs. chemistry,  $F(1,164) = 14.50, p < .001$ ; molecular biology vs. political science,  $F(1,154) = 66.10, p < .001$ ; and physics vs. cognitive psychology,  $F(1,164) = 21.47, p < .001$ ). The evolutionary biology vs. economics drop was in the right direction but did not reach significance,  $F(1,154) = 1.93, p = .17$ . There was also

a main effect of grade,  $F(2,149) = 4.37$ ,  $p < .02$ ,  $\eta^2 = .055$ . The grade by study interaction,  $F(2,149) = 4.79$ ,  $p < .01$ ,  $\eta^2 = .049$ , suggests that the difference between the Study 1 and Study 3 scores became larger with increasing age (sixth graders excluded).

### 4.3. Discussion

Two patterns emerge from Study 3. First, evidence was found for use of the schemas early on as the ability to cluster by discipline was clearly reduced in cases where discipline knowledge clashed with schemas that might be used to identify that discipline. Second, with increasing age, the ability to cluster by discipline was influenced less and less by the schemas used in this study. Transcripts from earlier studies suggest that children are using several schemas at the same time to help decide how to cluster knowledge and that pitting any one schema against a discipline will only cause a partial drop, which is why performance was not suppressed completely. It is difficult to know if the best possible schema was used in each case and whether the +schema/–discipline items should be seen as fully matched opposites of the –schema/+discipline cases, but the clear influence of the manipulation supports the general claim that mental structures roughly analogous to the schemas are at work in making discipline-based clusters early in development. The influence of the schemas presumably declined with increasing age because older children developed richer schemas that were more closely connected to the causal and relational patterns that underlie disciplinary clusters.

## 5. Study 4: Picture matching control (PMC)

It is important to make sure that the pictures that accompany the stimuli were only general memory support cues and did not provide on their own any additional cueing information about the disciplinary clusters. To check for that possibility a study was conducted with 33 children and 38 college students who were asked which of two pictures went better with an original example. The triads presented were the same as those used in Study 1.

### 5.1. Method

#### 5.1.1. Participants

Thirty-eight college undergraduates along with 17 kindergartners and 16 fourth graders (mean ages 5 years 7 months; 9 years 5 months) participated in this study. All were from the greater New Haven area and reflected the general demographics of that population.

#### 5.1.2. Materials and procedure

Each participant was presented with 16 sets of three pictures. The pictures were all hand drawn and corresponded to one of the everyday phenomena. The pictures were the same as those used in Study 1.

The three pictures were all presented at the same time, the first one on the top of the page, and the other two next to each other below on the same page. Participants were asked to decide

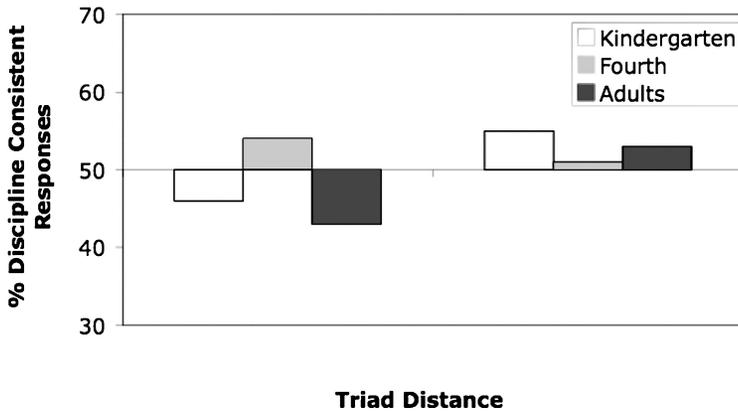


Fig. 6. Percentage of discipline-consistent answers made by children and adults on the picture-matching control study. Responses at all ages tended to cluster around chance levels of responding.

which of the two pictures on the bottom corresponded to the picture on top. Participants were not given any specific instructions about how to choose which pictures go best together but were asked to answer according to their intuitions.

## 5.2. Results

If participants chose the picture that corresponded to the discipline-consistent answer their responses were coded “1” and if they chose the picture that corresponded to the discipline-inconsistent answer their responses were coded “0.” The mean response patterns are shown in Fig. 6.

For children of both ages, all items were at chance levels ( $p > .05$ ). For adults, all items were at or below chance levels except for the molecular biology/political science and evolutionary biology/economics contrasts. This raises the possibility that adults may have done better on the economics/evolutionary biology contrasts because they were relying on some compatibility conveyed by the pictures alone.

Two arguments work against this possibility. First, if such differences were a factor, then those pictures that clustered at below chance levels should show a greater tendency to go in the opposite direction. A Pearson correlation was done to determine whether the pictures were in fact influencing participant’s decisions on the triad task. The percentage of adults who answered each triad correctly was correlated with the percentage of adults who chose the discipline-consistent picture on the PMC study. The results ( $N = 32$ ,  $r = -.017$ ,  $p = .928$ ) indicate that there is no correlation between the scores on the triad task and the PMC task and therefore they are unlikely to be influencing decisions on the triad task. Second, in the reverse triad condition, no significant correlations were found in a follow up analysis ( $N = 20$ ,  $r = -.221$ ,  $p = .599$ ) yet the same discipline clustering occurred. A similar pattern was found with the child data from the original triad and the PMC study ( $N = 16$ ,  $r = .159$ ,  $p = .556$ ).

### 5.3. Discussion

Study 4 rules out the possibility that participants in any age group were responding not on the basis of the verbal descriptions of who knows what but rather on the basis of which visual depictions they thought went together better. The pictures are extremely helpful to participants as a form of mnemonic support for the verbal descriptions, but on their own, they do not bias participants toward one choice over the other.

## 6. Study 5: Latent semantic analysis

Even the youngest children in the studies described in this article seem to have some insight into discipline-based ways of clustering knowledge. There are two abstract and conceptual ways that the children might be making these judgments. First, they might be sensing causal schemas that are particularly important in a broad domain, such as that, in physical mechanics, input forces tend to cause consequences in a manner that monotonically varies with the power of those inputs. Second, they might be sensing higher order categories that are linked to a domain, such as bounded objects in motion for mechanics or the presence of emotions for social psychology. Thus, the children who grouped two social psychology phenomena because they both had to do with feelings or who grouped two physics cases because the both had to do with gravity could have been doing so because they were referring to causal schemas of interpersonal and object interactions or because they identified a relational category into which both pairs fit.

These two strategies are difficult to distinguish on the basis of the data gathered so far because the children's explanations of their judgments are quite rare and because only some of those explanations explicitly mention causal relations. They could be implicitly invoking causal schemas when they mention an idea such as feelings, meaning that protagonist's feelings are causally influencing their actions in ways not seen for the other non-social psychological events in which protagonists are described; but they may not consistently be doing so. They might also be using causal notions in some cases, such as mechanics, and acausal relations in others, such as social psychology. Whichever of these strategies turns out to be dominant, however, they both need to be distinguished from another possible way of clustering knowledge that is based on more purely associative grounds. This alternative is suggested by a technique of modeling of associative information known as latent semantic analysis (LSA).

To what extent might the associative relations between words allow children to succeed in the clustering of knowledge as opposed to using schemas of the sort suggested in Study 3? To examine this possibility, a study was conducted using a program that ranks the conceptual similarity of word strings based on the co-occurrence patterns of words in those strings.

LSA is a process that extracts the contextual-usage meaning of words through statistical computations that are applied to a large corpus of text (Landauer, Foltz, & Laham, 1998). The theory driving the method is that the sum of the information about the contexts in which a particular word does and does not appear gives a set of mutual constraints that establishes the similarity of meanings of words and sets of words to each other. The similarity is not simply based on the frequency and co-occurrences of words but on higher order correlations

among correlations. Thus, two words are coded as semantically related not only if they co-occur with each other in a paragraph-sized chunk of text, they are also coded as related if they never directly co-occur but each co-occur with another common word. If, for example, “magnet” and “battery” rarely co-occurred in discourse, but if both frequently co-occurred with “electricity,” they would be coded by LSA as much more closely related than one would guess on the basis of the simple first-order co-occurrence of “magnet” and “battery.” During the stimulus construction phase for all strings used in Studies 1–3, attempts were made to use synonyms and paraphrases for words that might otherwise co-occur across more closely related items. LSA analyses are much more sensitive than simple surface co-occurrence, but this manner of stimulus construction may have helped reduce such statistical co-occurrence cues.

In LSA, the mathematical representation of a corpus of text is a semantic space that captures the multiple levels of co-occurrence. Within the semantic space, each word and combination of words, including novel combinations, has a multidimensional vector representation. When measuring the similarity between two words in the semantic space, one therefore measures the cosine of the angle between the vectors of the words (Landauer et al., 1998). The scores are standardized to range from  $-1$  to  $1$ .

In Study 5, we used semantic spaces corresponding to a third-grade reading level corpus, comprised of texts from various academic and other domains that amount to a total of 6,974 documents and a college-level reading corpus consisting of 37,651 documents. These different corpora are meant to represent the statistical patterns of word relations that children and adults at these levels would normally encounter. No kindergarten corpus was easily available for use with the LSA program, but if there are no indications of LSA relations cueing somewhat older children and adults, it is highly implausible that relations should suddenly be more salient with an even smaller corpus of the sort that would be associated with kindergarteners.

The triads used in Studies 1, 2, and 3 were submitted to the LSA engine. All words in each sentence after the word “why” were submitted. For example, in the sentence “Adam knows all about why people fight more when they are tired,” “people fight more when they are tired” would be submitted. All conceptual distances between the initial phenomenon sentence and the two choices were computed. The difference between the two LSA scores was computed; for example, for a chemistry/physics triad, the LSA distance score for the initial chemistry fact and the physics fact was subtracted from the LSA distance score for the initial chemistry fact and the other chemistry fact. The distribution of the difference scores was normal. A Pearson correlation was calculated to determine whether there was any correlation between the difference scores and the percent of adults who answered the triads according to discipline. The correlations were distributed normally around a value of approximately 0. Average values were then computed for all correlations used in each study with each grade corpus. Those values are shown in Table 5.

If a pattern of co-occurrence between words used in the different items was sufficient to provide information about how to cluster by discipline we would expect a strong correlation between the difference scores for the individual items and the percentage of participants who selected the correct discipline for each item. Note that in all cells in Table 5, the average correlations are very close to 0 and are far from significant (all  $p$  values  $>.05$ ). In

Table 5  
Average correlations between LSA distance scores between pairs of sentences and the extent to which those sentences were clustered together in the knowledge clustering task

	Study 1	Study 2	Study 3
3 <sup>rd</sup> grade	.026	.013	-.019
College corpus	.015	-.057	.056

addition, the correlations for the near items were not significantly different from the far items.

### 6.1. Discussion

Co-occurrence patterns are apparently not an available cue used in Studies 1, 2, and 3 as a way of clustering knowledge in the minds of others. This finding supports the argument that the children were not using patterns of association to make their judgments. We would not be surprised if adding schema-consistent associative information to stimuli helped performance. There is no reason why patterns of co-occurrences between words could not provide a separate useful heuristic for clustering knowledge. The key point of Study 5 is that the information was not available to the participants in our studies and is not necessary for the performance observed in adults and children.

## 7. Mechanisms of change

Although kindergartners have some sense of discipline-based ways of clustering knowledge, that sense seems relatively fragile in comparison to older children and adults. Thus, although the discipline-based manner of clustering is available for contrasts between relatively distant or far domains, it is not dominant or especially salient to younger children. In addition, as mentioned earlier, when discipline-based ways of understanding the division of cognitive labor are pitted against clustering knowledge by goals and surface topic, younger children are very often drawn to the goal- and topic-based ways of clustering (Danovitch & Keil, 2004). What sorts of factors might help the developing child see more clearly the special status of discipline-based forms of knowledge? Studies 6 and 7 explore two possible ways in which children might come to better appreciate the special significance of discipline-based forms of knowledge. Both studies are organized around the idea that certain contexts may help highlight the special status of disciplinary-based knowledge. We can then ask if children are sensitive to such contexts and their links to disciplinary knowledge and whether that sensitivity is early emerging or comes in only later in childhood. We predicted that even young children might benefit from such contexts and that, over time, those contexts might help highlight the special status of discipline-based forms of expertise.

## 8. Study 6: Levels of categorization and forms of expertise

One clue to the special status of discipline-based knowledge as opposed to other forms may reside in the understanding that higher levels of knowledge categorization become decreasingly plausible when understood in topic-based or goal-based ways. If Martin is described as knowing a great deal about Lhasa Apsos, it is plausible that he is a kind of Lhasa Apso fanatic who knows almost everything there to know about that breed of dog. Martin would know some distinctive aspects of Lhasa biology but would also know about the history of Lhasa Apsos in Tibet, how they were smuggled out of country by the British, what films have had Lhasas in them, how much china figurines of Lhasas normally cost on eBay, and so on. Higher up in the hierarchy of kinds, however, this sort of trivia nut knowledge becomes less plausible. One cannot be an animal expert by knowing everything that is associated with animals—the category is far too vast. As one moves up hierarchies of categories, one needs a way to limit the knowledge about the category to a meaningful and coherent set of information. Discipline-based understanding works very well in that respect. An expert on animals would therefore be expected to know general principles of biology that apply to all animals and not trivia that is relevant to various species and subspecies, as generic animal trivia is very rare. Discipline-based knowledge is one of the few plausible ways of thinking about expertise for categories above the basic level of categorization (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976).

It may therefore be that as children are exposed to more and more high-level category words, or to discourse and problems concerning high-level categories, they start to realize more clearly the special status of discipline-based knowledge. Indeed, this process may be similar to the ways some kinds of words are seen by young children as “invitations to form categories” (Waxman & Markow, 1995, p. 298) and the ways in which certain generics prompt essentialist thinking (Gelman & Raman, 2003). The high-level labels are not the basis of discipline-based knowledge but might well help to elevate it to a more central and privileged way of organizing knowledge. In most cases, terms above the basic level of categorization are less frequent and acquired later in development (Mervis, 1987) and so it is plausible that the elementary school years are a period during which the increasing use of high-level category names helps shift children toward favoring a more robust discipline-based way of viewing knowledge. If higher level category names have that effect, it should be possible to manipulate the extent to which discipline-based clusters are chosen by showing that the use of higher level names results in more discipline-based judgments. Study 6 explored this possibility.

### 8.1. Method

#### 8.1.1. Participants

Forty-eight children participated in this study: 16 kindergarteners, 16 second graders, and 16 fourth graders (mean ages, 5 years 3 months; 7 years 2 months; 9 years 2 months). Children were interviewed individually outside of their classrooms, with each session lasting approximately 20 minutes. There were roughly equal numbers of males and females in each

condition. All were from the greater New Haven area and reflected the general demographics of that population.

### 8.1.2. Materials and procedure

8.1.2.1. *Stimuli*: The stimuli were constructed using a similar format to that used in Studies 1–3. A person was described as knowing “a lot” about either a high-level category (animals, plants, metals) or a low-level category (dolphins, apples, gold). Participants were then asked to judge which of two facts concerning the low-level category the person would know more about: a fact that is highly associated with the category but has no relation to the discipline (i.e., trivia) or a fact that is discipline centered. For example:

Jeff knows a lot about animals (dolphins). Which would he know more about?

1. Why dolphins are in an exhibit in Sea World?
2. Why dolphins need to have hearts to keep them alive?

For each case, half of the participants received the high-level category (animals) and half received the low-level category (dolphins). All participants received a counterbalanced mixture of the high- and low-level category labels for each of three types of stimuli: animals, plants, and metals (see Appendix E for full list of stimuli). A person who is an expert in the low-level category might know either biological facts or trivia, but an expert in the high-level category would be more likely to know biological facts than trivia. Hence, the main pattern of interest was whether participants shifted more toward discipline-based responses for high-level categories. We did not expect judgments about the expert at the low-level category to be completely trivia based, just more so than those about the expert whose initial knowledge is said to be at a higher level.

8.1.2.2. *Procedure*: The procedure was the same as that used in Studies 1–3. Based on the first fact, children were asked which of two phenomena the expert was more likely to know about. The same training items used in Study 1 and 2 were used in this study.

## 8.2. Results

Each child was assigned a score of “1” for discipline-based answers and “0” for trivia-based answers and an average score was then calculated. Our initial analyses revealed no category type (animals, plants, metals) by level interaction,  $F(2,90) = .100$ ,  $p = .905$ ,  $\eta^2 = .002$ , so we collapsed across the category types for the remainder of the analyses. Using a  $2 \times 3$  repeated measures ANOVA with category level (high vs. low) as a within-subject variable and age (K, 2, 4) as a between-subject variable, we analyzed the degree to which the category level influenced participants’ discipline-based responses.

Figure 7 shows the degree to which discipline clusters were preferred for high-level categories as opposed to low-level categories. There was a main effect of category level with more clustering by discipline with the superordinate category labels ( $M = 86.43\%$ ) than basic level category labels ( $M = 53.75\%$ ),  $F(1,45) = 48.77$ ,  $p < .001$ ,  $\eta^2 = .520$ . There was no

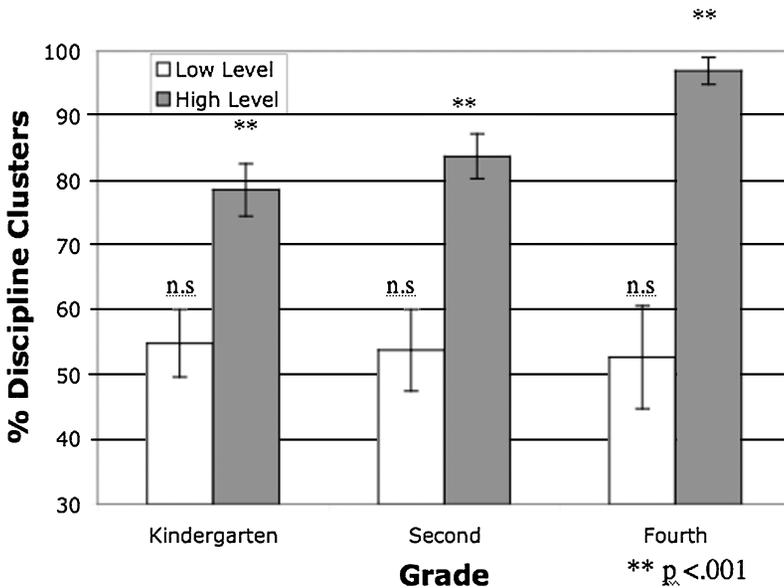


Fig. 7. Children at all ages were more inclined to pick a discipline-based way of clustering knowledge when higher order category labels are used.

main effect of grade,  $F(2,45) = 1.66$ ,  $p = .201$ ,  $\eta^2 = .069$ , and no level by grade interaction,  $F(2,45) = 1.69$ ,  $p = .197$ ,  $\eta^2 = .070$ .

### 8.3. Discussion

Children of all ages in this study grouped knowledge in a more disciplinary-based manner when the expert's knowledge was described at a high level of categorization. If the initially attributed knowledge was described with a superordinate level term, such as "animal," "plant," or "metal," the children were much more likely to pick the type of knowledge in the subsequent pair that was in the same discipline. By contrast, if the initially attributed knowledge was described with a basic or subordinate-level term such as "poodle," "pumpkin," or "silver," children tended to choose evenly between discipline-based and trivia-based ways of grouping. Use of a low-level term did not rule out discipline-based ways of grouping, it merely made trivia-based ways much more plausible and therefore usually resulted in roughly equal choices between the two ways of grouping.

Even the youngest age group was differentially influenced toward discipline-based clusters when high-level categories were used in the initial expertise item. Indeed, the effect of discipline-based categories was equally strong at all ages. Thus, at least as early as 5 years of age, the presence of higher level category labels causes a strong shift toward discipline-based groupings of knowledge. This is, of course, only the first step in inferring that, in development, higher level categories actually cause a shift toward favoring discipline-based groupings. It may be, for example, that children show increased use of higher level category labels with increasing age just because they are gaining disciplinary insight that makes those labels more

meaningful, rather than the other way around. The direction of causality here might be explored in future studies that manipulated over the course of an extended period, such as a school year, exposure to higher and lower level category terms and looked at the effects of knowledge clustering with other terms.

## 9. Study 7: Relations versus category labels

There is now an extensive body of research exploring how patterns of induction change during preschool and the elementary school years (Carey, 1985; Gelman, 2003; Inagaki & Hatano, 2002). That work consistently demonstrates an emerging ability to make inductions over abstract kinds such as living things and an accompanying ability to vary one's inductions based on the kinds of properties involved; e.g., whether properties are biological or not. In addition, judgments of particular biological processes such as inheritance show a growing appreciation that certain kinds of properties, such as internal physical structure, are more likely to be linked to those biological processes than others, such as social relationships (Inagaki & Hatano, 2002; Springer & Keil, 1991). Thus, advances in biological knowledge during the elementary school years are associated with more fine-grained intuitions about higher level biological categories and about biological properties. As this sort of knowledge develops and as children see the same entities embedded in a wider range of contexts that suggest different causal/relational patterns, those contexts may help children more clearly see how discipline-based clusters of knowledge vary not by the entity involved per se but in terms of how specific kinds of properties are related to particular causal and relational patterns. Thus, an appreciation of discipline-based knowledge may be strongly related to an awareness that such knowledge is, at the deepest level, organized around distinct clusters of specific property types and their causal/relational patterns, not just around surface objects that can be proxies for several different ways of construing expertise.

The most systematic studies of interactions between property types and the different causal relations underlying a kind have been done in adults (e.g., Heit & Rubinstein, 1994). For example, if taught an "anatomical" property (e.g., "all its cells contain a small amount of zinc") on one animal, participants are more likely to inductively extend it to an animal of the same broad taxonomic group (e.g., whales and bears as mammals) than to an animal showing similar behaviors (e.g., whales and tuna as swimming creatures). If, however, participants are taught a behavioral property on the same animal (e.g., moves in zig-zag manner), they will extend it more to behaviorally similar animals than taxonomically similar ones (Heit & Rubinstein, 1994). We see this ability to shift patterns of induction as a function of property types as an indication of an appreciation of the centrality of causal relational patterns to thinking about categories and not just the objects that comprise the categories. If children can show a similar ability in making inductions about knowledge states, that would support the view that children use deeper causal processes to cluster knowledge.

Emerging disciplinary understanding may therefore be linked to an ability to step back from thinking about object-centered expertise to thinking more about expertise-based on relations between property types. Thus, a discipline-based expert is not an expert on all aspects of elephants or zebras but only on biologically related properties of zebras and elephants.

Since there is considerable improvement in the discipline-based judgments between ages 5 and 10, perhaps children also show a corresponding improvement in understanding how property by entity relations are more critical to discerning disciplinary knowledge than a mere entity label. If such linkage could be found, it might shed insight on how discipline-based clusterings of knowledge emerge. That is, as a child becomes more sophisticated in disciplinary understanding, she might be able to realize the special importance of focusing on property types and their relations rather than mere objects and use that insight as a general strategy for understanding the division of cognitive labor. A study was therefore designed in which the relational structure in a domain contrasted with the topic of study. Consider the following example:

Danny knows all about what dogs and kittens are **made up of**  
 Does he know more about what foxes or what parrots **are made up of**?

Vs.

Danny knows all about **where to buy** dogs and kittens  
 Does he know more about **where to buy** foxes or parrots?

In the first example, most adults judge that Danny knows more about what foxes are made up of because the inside parts of animals invokes a biological realm and consequently a clustering of animals on the basis of their biological similarity. By contrast, in the second example, most adults judge that Danny knows more about where to buy parrots, because the best place to purchase animals suggests a different set of relations having to do with the economics of pet stores. Adults, at least, do not cluster knowledge in the mind of others merely on the basis of the most prominent kinds that are mentioned. Both pairs of sentences describe dogs, kittens, foxes, and parrots. They are sorted differently, however, as a function of being understood in biological vs. economic terms. Thus, discipline-based pairing might vary dramatically depending on the relational structures involved. This study explores how such relation-based similarities might emerge in development and how children learn to distance themselves from objects per se and focus more on relations between property types in making judgments about the division of cognitive labor.

## 9.1. Method

### 9.1.1. Participants

Seventy-two children participated in this study: 24 each in kindergarten, second, and fourth grade (mean ages 6 years, 0 months; 8 years 1 month; 9 years, 10 months). Children were interviewed individually outside of their classrooms with each session lasting approximately 20 minutes. There were roughly equal numbers of males and females in each condition. All were from the greater New Haven areas and reflected the general demographics of that population.

### 9.1.2. Materials and procedure

9.1.2.1. *Stimuli*: The stimuli were constructed using the format described in the example above. A person was said to know all about the best place to buy an animal (invoking

economic relations) or what the animal was made up of (invoking biological relations) and were then asked which of two other animals they would know more about: a biologically similar animal or an economically similar animal (see Appendix F for full list of stimuli). Half of the 14 pairs included animals that are typically considered pets (e.g., dogs, kittens, parrots) and half included animals typically found on a farm (e.g., horses, sheep, cows). Each participant received 7 of the pairs described in the economics invoking contexts (where to buy) and 7 described in the biology invoking contexts (made up of). The distribution of the stimuli were counterbalanced such that the animals used in the biology-invoking context for half of the participants were then used in the economics-invoking context for the other half of the participants. The order of presentation was randomized.

*9.1.2.2. Procedure:* All children were given two training items to acquaint them with the format of the task. No children needed to be excluded for missing the training items. The procedure was similar to that used in Studies 1–3 such that each participant was told about a person who knows all about something and asked to choose which of two things they would also know more about.

## 9.2. Results

If participants chose the biology answer, their responses were coded as “1” and if they chose the economics answer their responses were coded as “0.” A total score was calculated for the different contexts (biology and economics). Mean scores are shown in Fig. 8. Our initial analyses revealed no stimulus type (farm animal or pet) by context interaction,  $F(1,69) = .66$ ,  $p = .42$ ,  $\eta^2 = .003$ , and therefore we have collapsed across stimulus type for the remainder of our analyses. A  $2 \times 3$  repeated measures ANOVA with context (economics vs.

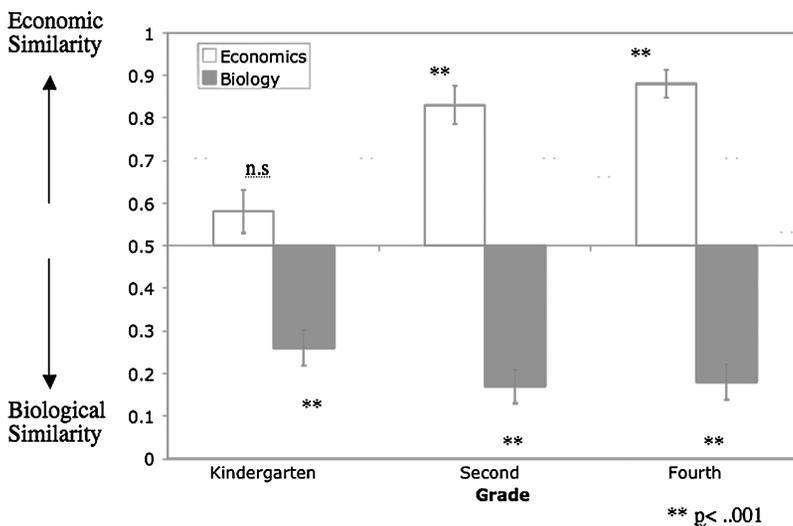


Fig. 8. Scores above .5 represent a choice based on economic similarity and those below represent a choice based on biological similarity.

biology) as the within-subjects variable and age (K, 2, 4) as the between-subjects variable was run to determine the effect of context on participants' groupings. There was a significant main effect of context,  $F(1,69) = 254.84$ ,  $p < .001$ ,  $\eta^2 = .787$ , suggesting that participants were highly influenced by the information surrounding the decision (economics  $M = .77$ ; biology  $M = .20$ ).

There was also a main effect of grade,  $F(2,69) = 3.54$ ,  $p < .05$ ,  $\eta^2 = .093$ , suggesting an increase in the correct answer choice with age. Even the youngest children in this task respond differently in the two contexts, clearly favoring the biological similarities in the "made up of" contexts and choosing more equally between alternatives in the "where to buy," as can be seen by the context by grade interaction,  $F(2,69) = 11.55$ ,  $p < .001$ ,  $\eta^2 = .251$ . There was also a suggestive change between the kindergarten and second-grade children in the ability to differentiate these two kinds of contexts with a larger portion of the change occurring with the economics contexts, but it did not reach significance.

### 9.3. Discussion

If children cluster knowledge in the minds of others based on a global assessment of the similarities between the most prominent objects discussed in examples, they should not differentiate between objects embedded in different contexts. For example, dogs, kittens, foxes, and parrots can only be similar to each other in one global way. Yet, even the youngest children in this study showed a sensitivity to the centrality of property types and their relations instead of merely the object of study. Thus, early on, they start to realize that expertise is often more valuably organized around key relational patterns than it is about everything that is associated with members of a category. This awareness may grow stronger as they see different sets of relational patterns applied to the same set of entities and realize that the relational patterns are more useful to clustering many objects than using the entity names on their own. Children start to realize that expertise is not so much about a basic-level object as about a more relational construct (e.g., buying and selling or how essences cause surface properties). This pattern is likely related to other studies showing developmental shifts in which there is an increasing emphasis on relational forms of similarity (Gentner & Namy, 1999; Markman & Gentner, 2000). It may therefore be that increased exposure to different kinds of properties and their links to causal and relational patterns helps children to see the special richness and power of those patterns that are related to disciplinary thought.

## 10. General discussion

Children as young as 5 years of age show the beginnings of an ability to infer clusters of knowledge in other minds in a manner that crudely maps on to the major academic divisions of the natural and social sciences. In Study 1 even children as young as 5 sorted three out of four pairs of domains on a disciplinary basis and, in Study 2, they sorted one out of four such domains. By second grade, children were sorting almost pairs of domains correctly in both studies, suggesting that they tap into the causal and relational structures of both domains that make up the triads. When children succeed, they do so without exposure to those

academic disciplines and with apparently little or no understanding of the details of those disciplines. They seem to accomplish this sorting by referring to simple schemas for causal patterns distinctive to domains, schemas that can lead them astray when a phenomenon is in the discipline but lacks the superficial signs associated with the schemas. This pattern was demonstrated in Study 3 where younger children were especially dependent on simpler and more misleading schemas. It is clear, however, that even the youngest children do not need to use shallower tricks to engage in discipline-based clusterings. Study 4 showed that there were no clues in the stimulus pictures by themselves and Study 5 showed that the children were succeeding in clustering knowledge even when associative patterns were noninformative as to areas of expertise.

Studies 6 and 7 explored two other abilities that might be related to disciplinary insight: the realization that discipline-based expertise is more plausible than other forms at higher levels of categorization (Study 6) and the understanding that expertise clusters are based on interactions of particular property types and certain causal and relational patterns (Study 7). Even 5-year-olds showed both abilities, with a significant improvement in the ability with age in Study 7.

The discipline-based way of construing the terrain of knowledge, however, is quite fragile early on and is in tension with other ways such as through goals or surface topics. By the time children are around 10 years of age, the discipline-based way of construing knowledge is considerably more robust and has a more privileged status as a way of understanding how knowledge is organized and distributed across other minds. Two questions arise from these results: (a) How do the children, and for that matter adults, succeed in clustering knowledge on the basis of academic disciplines when they know so little about those disciplines? (b) What changes during the elementary school years to help make the discipline-based way of construing knowledge emerge as a more powerful default option?

With respect to the first question, two interrelated strategies may be used. First, children might use intuitive skeletal schemas of the kinds of relations that are central to a domain, apparently operating under the assumption that phenomena covered by the same schemas are likely to arise from the same deep principles. These skeletal schemas are likely to be closely related to what are known as core or framework theories in young children. A person who fully understands any one of the phenomena in a cluster would therefore be very likely to understand others because he knows the common principles underlying and explaining all of them. Second, we can cluster phenomena by their membership in higher level abstract categories. Thus, we might cluster a group of phenomena as social psychology because they all involve social interactions or a group as evolutionary biology because they all involved adaptations to niches.

How different are these strategies? To what extent can one pick out an abstract category without having the relevant schemas? In many cases, use of a high-level abstract category label may only be effective when the label is grounded in a sense of the relational patterns that make up that category. Thus, a child might cluster together a group of phenomena in the same domain of physics expertise because they all involved dynamic, causal interactions between bounded objects, such as launching events. Alternatively, a child might cluster merely on the basis of the presence of bounded objects in motion. If this second strategy required no schematic notions of relations (causal or otherwise) between objects, then it would appear to

be a distinct strategy. In other domains, an abstract category such as “feelings” may be more intrinsically relational (i.e., anger presupposes an agent in a specific mental relation to another entity) and therefore harder to distinguish from schematic knowledge.

A related question concerns whether a sense of specifically causal patterns is used to cluster knowledge. Such an account requires a level of representation largely devoid of detailed mechanistic knowledge, an account that is compatible with recent work suggesting the presence of such abstract causal schemas in the first year of life (Mandler, 2004). Use of causal schemas, however, does not mean that children, and adults, might not be able to sort expertise in abstract relational ways that do not involve notions of cause. Thus, a sense of the abstract relational patterns of mathematics or language might also allow them to cluster knowledge corresponding to those areas of expertise; and more broadly, even in those domains where causal patterns are more clearly at work, the causal relations may not always need to be grasped. We suspect that causal relations may play a central role through the developmental period studied here, but that supposition needs further support in which both noncausal relational patterns and nonrelational abstract categories are considered as alternatives. More elaborate versions of Study 3, which pit disciplines against different sorts of schema and category cues, might reveal which sorts of categories are used as well as whether there is a developmental shift in the basis used for clustering. More broadly, it appears that at about age 5, children begin to tap into their intuitive understandings in distinct domains that they inherit from infancy and their preschool years. It is also unlikely that young children are simply clustering the phenomena together without thinking about how those phenomena are related to areas of expertise. Ongoing studies, for example, show that embedding the phenomena in contexts such as “John knows all about why x” causes adults to focus on discipline-based clusters much more strongly than when the bare phenomena are presented and other dimensions of similarity are more salient.

The ability to discern disciplinary clusters of knowledge developed considerably during the ages studied. Younger children had considerable difficulty in making finer distinctions among the disciplines, such as between social and cognitive psychology. As children learn ever more subtle patterns of regularities, their growing theories of the world thus influence their theories of the division of cognitive labor. The process is a lifelong one. Even sophisticated adults can be shown to switch in how they cluster knowledge as their insights into deeper principles develop (Chi, Feltovich, & Glaser, 1981).

Because discipline-based clusters are only one of several possible ways to cluster knowledge in other minds, part of what develops seems to be a sense that discipline-based forms of knowledge have a special privileged status in terms of their generative explanatory power. Studies 6 and 7 explored two possible mechanisms that might be involved in a growing appreciation of the special status of discipline-based forms of knowledge. One involves an increasing use of superordinate domains of expertise, where the superordinate domains are difficult to understand in any other way than discipline based. We know from studies of word usage and vocabulary growth that the use of high-level category terms increases during the elementary school years (Mervis, 1987). Increased exposure to terms for high-level categories may foster an awareness of the distinct advantage that discipline-based clusters offer for understanding knowledge at those levels. This may be part of a broader set of strategies in which children see words as invitations to form categories (Waxman & Markow, 1995). Alternatively, increasing disciplinary knowledge in a domain may make higher level categories

more meaningful and useful. The direction of causal influence here might be uncovered by manipulations of exposure to higher level categories and seeing whether increased use of higher level category labels caused increased disciplinary sorts later in tasks that no longer had the explicit support of those labels.

The second possible mechanism related to an increasing awareness of the special nature of discipline-based knowledge may be seen in the ways in which growing knowledge in a domain fosters an awareness of the importance of property types and relations over objects as most relevant to expertise (Gentner & Namy, 1999). In Study 7 an economic frame and a biological frame resulted in two contrasting ways of sorting expertise. Again, there were clear signs of a sensitivity to such frames in even the youngest children studied. As children grow older, they start to distance themselves more and more from object-centered forms of expertise and consider instead how objects enter into patterns of relations with distinct sets of property types. Understanding such relations is at the core of appreciating the special status of discipline-based knowledge. Thinking of people as dog or kitten experts is much less helpful than thinking of people as experts on how inside parts work together to explain surface properties or as experts on how cost and value are determined. Future studies will have to determine whether there is a relatively domain-general emerging appreciation of the importance of relations between property types over objects or whether this appreciation is closely linked to developing insights in each domain.

The child's developing understanding of the world often seems to presuppose a "stubborn autodidact" model (Harris, 2001, p. 495), in which the child gains all knowledge through direct experience with the world. It is easy to visualize children as lone individuals observing phenomena and constructing all their beliefs and understandings based on those real-world observations. This perspective, however, cannot possibly be true. Vast amounts of knowledge are gathered through accessing information in other minds. Virtually all of history must be acquired in this way as well as knowledge about microscopic or distant aspects of the world. A critical part of that process is learning what experts to approach for what kinds of knowledge. Having a sense of the division of cognitive labor provides an essential tool for gathering further knowledge and knowing how to rely on expertise in other minds. Although young children may sometimes be more accepting of adult claims than adults, they are also aware of different kinds of experts and that people's claims may not always be correct (Mills & Keil, 2005). More broadly, it has long been known that young children are aware that other individuals have different knowledge bases that they bring to the conversational arena (Shatz & Gelman, 1973).

In a classic paper published in 1975 entitled "The Meaning of Meaning," the philosopher Hilary Putnam (1975) argued that our referential practices with words are critically supported by a "division of linguistic labor" (p. 131) in which a person may successfully and appropriately use a word as long as he or she belongs to a linguistic community in which there are experts who can correctly pick out the word's referents. The person may not be able to pick out the referents but may leverage his crude understanding of the word's meaning through potential access to the right experts. This argument makes clear just how much of the meaning of words depends on the broader communities in which we live and on a sense of how those communities are organized. The division of linguistic labor is part of a much broader phenomenon known as the *division of cognitive labor*, where the reliance on the contents of other minds is equally,

if not more, important. It is a hallmark of all cultures, traditional or technological, to have both a division of cognitive labor and a means of navigating those divisions.

David Hume stated in vivid terms the advantages of the division of labor:

When every individual person labours a-part, and only for himself, his force is too small to execute any considerable work; his labour being employed in supplying all his different necessities, he never attains a perfection in any particular art; and as his force and success are not at all times equal, the least failure in either of these particulars must be attended with inevitable ruin and misery. Society provides a remedy for these three inconveniences. By the conjunction of forces, our power is augmented: By the partition of employments, our ability encreases: And by mutual succour we are less exposed to fortune and accidents. It is by this additional force, ability, and security, that society becomes advantageous. (Hume, 1739/2000)

If we are to similarly benefit from the division of cognitive labor and thereby provide mutual succor to each other's minds, we cannot simply each rely on our own distinct bodies of knowledge. We must have some sense of who knows what in the world around us and how to access that information as well. One of the most powerful ways of thinking about the organization of knowledge is based on the following idea: There are domain-specific patterns in the world that experts know and use to understand a wide range of phenomena that arise from those patterns. If one can also grasp those patterns in some coarse yet reliable manner, one knows which sort of expert to approach for further understanding. We have shown here that this appreciation of discipline-based ways of understanding shows its first signs quite early in childhood and develops substantially during the elementary school years.

## Acknowledgments

This research was supported in part by NIH Grant R37-HD23922 to Frank Keil. We are grateful to the staff, parents, and students of St. Margaret's—McTernan School, Kindercare Early Learning Center, Clintonville Elementary School, Long Meadow Elementary School, the Southington Public School District, and the many members of the greater New Haven community. We also thank members of the Yale Cognition and Development Laboratory for their insightful comments; Paul Bloom, Kristi Lockhart, Art Markman, and three anonymous reviewers for extensive comments on drafts of this article; and Azi Fathi for help in the execution and design of two studies.

## References

- Au, T. K. (1994). Developing an intuitive understanding of substance kinds. *Cognitive Psychology*, 27, 71–111.
- Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, 11, 211–227.
- Bruner, J. (1960). *Process of education*. Cambridge, MA: Harvard University Press.
- Carey, S. (1985). *Conceptual change in childhood*. Bradford Books: MIT Press.
- Casson, L. (2001). *Libraries in the ancient world*. New Haven, CT: Yale University Press.
- Chi, M. T. H., Feltovitch, P. J., & Glaser, R. (1981). Categorization of physics problems by experts and novices. *Cognitive Science*, 5, 121–152.

- Chinn, C. A., & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. *Review of Educational Research*, 63, 1–49.
- Culler, J. (1976). *Structuralist poetics: Structuralism, linguistics, and the study of literature*. Ithaca, NY: Cornell University Press.
- Danovitch, J., & Keil, F. C. (2004). Should you ask a fisherman or a biologist? Developmental shifts in ways of clustering knowledge. *Child Development*, 75, 918–931.
- diSessa, A., Gillespie, N., & Esterly, J. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science*, 28, 843–900.
- Durkheim, E. (1997). *The division of labor in society* (G. Simpson, Trans.). New York: The Free Press. (Original work published 1893)
- Fricker, E. (2002). Trusting others in the sciences: A priori or empirical warrant? *Studies in the History and Philosophy of Science*, 33, 373–383.
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford: Oxford University Press.
- Gelman, S. A., & Raman, L. (2003). Preschool children use linguistic form class and pragmatic cues to interpret generics. *Child Development*, 74, 308–325.
- Gentner, D., & Namy, L. (1999). Comparison in the development of categories. *Cognitive Development*, 14, 487–513.
- Goldman, A. (2002). *Pathways to knowledge*. Oxford, England: Oxford University Press.
- Hacking, I. (1999). *The social construction of what?* Cambridge, MA: Harvard University Press.
- Harris, P. L. (2001). Thinking about the unknown. *Trends in Cognitive Science*, 5, 494–498.
- Hatano, G., & Inagaki, K. (1999). A developmental perspective on informal biology. In D. L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 321–354). Cambridge, MA: MIT Press.
- Heit, E., & Rubenstein, J. (1994). Similarity and property effects in inductive reasoning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 20(2), 411–422.
- Hume, D. A. (2000). *A treatise on human nature*. D. F. Norton & M. J. Norton (Eds.). New York: Oxford University Press. (Original work published 1739)
- Inagaki, K., & Hatano, G. (2002). *Young children's naive thinking about the biological world*. New York: Psychological Press.
- Keil, F. C. (1979). *Semantic and conceptual development: An ontological perspective*. Cambridge, MA: Harvard University Press.
- Keil, F. C. (1983). On the emergence of semantic and conceptual distinctions. *Journal of Experimental Psychology: General*, 112, 357–389.
- Keil, F. C. (2003). Folkscience: Coarse interpretations of a complex reality. *Trends in Cognitive Science*, 7, 368–373.
- Keil, F. C. (2006). Cognitive science and cognitive development. In W. Damon & R. Lerner (Series Eds.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol 2: Cognition, perception, and language* (6th ed., pp. 609–635). New York: Wiley.
- Kitcher, P. (1993). *The advancement of science. Science without legend, objectivity without illusions*. Oxford, England: Oxford University Press.
- Kitcher, P. (2001). *Science, truth and democracy*. Oxford, England: Oxford University Press.
- Krajcik, J. S., Czerniak, C. M., & Berger, C. (1999). *Teaching children science*. Boston: McGraw-Hill.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, 25, 259–284.
- Lawson, R. L. (2006). The science of cycology: Failures to understand how everyday objects work. *Memory and Cognition*, 34, 1667–1675.
- Lutz, D. J., & Keil, F. C. (2003). Early understandings of the division of cognitive labor. *Child Development*, 73, 1073–1084.
- Mandler, J. M. (2004). *The foundations of mind*. Oxford, England: Oxford University Press.
- Markman, A. B., & Gentner, D. (2000). Structure mapping in the comparison process. *American Journal of Psychology*, 113, 501–538.

- Medin, D. L., & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Artony (Eds.), *Similarity and analogical reasoning* (pp. 179–195). New York: Cambridge University Press.
- Mervis, C. B. (1987). Child-basic objects categories and early lexical development. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 201–233). Cambridge: Cambridge University Press.
- Mills, C. M., & Keil, F. C. (2005). The development of cynicism. *Psychological Science, 16*, 385–390.
- National Research Council. (2006). *Taking science to school: Learning and teaching science in grades K–8*. National Academies Press.
- Pelikan, J. (1992). *The idea of the university: A re-examination*. New Haven, CT: Yale University Press.
- Putnam, H. (1975). The meaning of “meaning.” In K. Gunderson (Ed.), *Language, mind, and knowledge* (Vol. 2, pp. 131–193). Minneapolis: University of Minnesota Press.
- Rosch, E., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 133–156.
- Rozenblit, L. R., & Keil, F. C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science, 26*, 521–562.
- Scholl, B. J., & Tremoulet, P. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4*, 299–309.
- Shatz, M., & Gelman, R. (1973). The development of communication skills: Modifications in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development, 38* (5, Serial No. 152).
- Simons, D. J., & Keil, F. C. (1995). An abstract to concrete shift in the development of biological thought: The insides story. *Cognition, 56*, 129–163.
- Smith, A. (1904). *An inquiry into the nature and causes of the wealth of nations* (5th ed.). London: Methuen & Co. (Original work published 1776)
- Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*, 605–632.
- Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 39–67). New York: Cambridge University Press.
- Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford, England: Blackwell.
- Sperber, D., & Hirschfeld, L. (2004). The cognitive foundations of cultural stability and diversity. *Trends in Cognitive Science, 8*, 40–46.
- Springer, K., & Keil, F. C. (1991). Early differentiation of causal mechanisms appropriate to biological and nonbiological kinds. *Child Development, 62*, 767–781.
- Ward, T. (2006). English law’s epistemology of expert testimony. *Journal of Law and Society, 33*, 572–595.
- Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology, 29*(3), 257–302.
- Wellman, H. M. (1992). *The child’s theory of mind*. Cambridge, MA: MIT Press.
- Wellman, H. M., & Gelman, S. A. (1998). Knowledge acquisition in foundational domains. In W. Damon (Series Ed.) & D. Kuhn & R. S. Siegler (Vol. Eds.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp. 523–530, 538–554). New York: Wiley.
- Werner, H., & Kaplan, B. (1963). *Symbol formation: An organismic-developmental approach to language and the expression of thought*. New York: Wiley.
- White, P. A. (2006). The role of activity in visual impressions of causality. *Acta Psychologica, 123*, 166–185.

Copyright of Cognitive Science is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.