

HANDBOOK OF CHILD PSYCHOLOGY

FIFTH EDITION

Volume 2: Cognition, Perception, and Language

Editor-in-Chief

WILLIAM DAMON

Volume Editors

DEANNA KUHN and ROBERT S. SIEGLER



John Wiley & Sons, Inc.

New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

CHAPTER 12

Enabling Constraints for Cognitive Development and Learning: Domain Specificity and Epigenesis

ROCHEL GELMAN and EARL M. WILLIAMS

KEY POSTULATES ABOUT LEARNING AND COGNITION 576

Postulate 1: Inputs Can Be Divergently and Convergently Ambiguous, Requiring Interpretation 577

Postulate 2: The Mind Is an Active Learner, Not a Passive Receptacle 579

Postulate 3: Cognition and Learning Are Always Selective 579

Postulate 4: Animals, Infants, and Young Children Are Not Perception-Bound Learners 581

Postulate 5: Learning Is Not a Single "Process" Since There Are Different Domain-Specific Kinds of Learning 583

Postulate 6: Epigenesis Leads to Variability and Requires Supporting Environments 584

Postulate 7: Environments and Minds Share Universal Structures, Not Universal Surface Characteristics 585

DIFFERING THEORETICAL PERSPECTIVES ON CONSTRAINTS AND LEARNING 591

About Terms and Their Role in Scientific Theories 592

Constraints in Associationist Theories 593

Constraints and Mental Structure I:

Information Processing Theories 594

Constraints and Mental Structure II: Stage Theories of Cognitive Development 596

Constraints and Mental Structure III: Domain-Specific Theories of Concept Learning 596

Summary 597

ENABLING CONSTRAINTS, EPIGENESIS, AND DOMAIN SPECIFICITY IN LEARNING 597

Constraints as Learning Enablers:

An Evolutionary Perspective 600

What Enabling Constraints Are and Are Not 602

Problems with Alternative Views of Constraints 604

Characterizing the Learning Environment 607

Learning Structured Information: The Constructivist

Learning Laws of Redundancy and Ubiquity 608

Sources of Structured Information in the

Learning Environment 609

Mental Learning Tools: Beyond Initial Enabling

Constraints and Core Domains 611

CONCLUSION 618

ACKNOWLEDGMENTS 619

REFERENCES 620

All theories of cognitive development make assumptions about constraints on learning, even those which commonly are seen as nonconstraint theories. We do not ask whether a constraint theory or a nonconstraint theory is better. Instead, we ask what kind of theory best accommodates key facts about cognitive development and concept learning. Our answer follows in three sections. We begin by reviewing what we consider these relevant facts to be, in the form of seven postulates about cognitive development and learning.

In the second section, we turn to a review of theories and the notions of constraint embedded within them. This allows us to bring out key sources of misunderstanding and confusion about what constraints are and how they function. We emphasize the fact that the meaning of theoretical terms including *constraint* and *learning* can differ considerably across theories. For example, some treat constraints as *learning enablers*, while others consider them to be limiting mechanisms that make learning unnecessary. We also point out instances in which seemingly different theories share overlapping assumptions about the nature of constraints.

In the third section, we move to a combined consideration of evidence and theory, and develop our reasons for

Partial support for the work came from NSF Grant No. DBS-92009741 to R.G.

favoring a rational-constructivist theory of cognitive development. The rationalist side of the theory is motivated in part by the many findings of conceptual competence in infants and animals, competencies that extend well past the simple perceptual abilities to which they were once thought to be limited. The constructivist side is motivated by the evidence that learning at all ages involves an active interchange between structures in the mind and "structure-able" information from the environment. We share with Piaget the hypothesis that our young join their mental structures with processes, such as assimilation and accommodation, to actively contribute to their own cognitive development and epigenesis.

Rather than pairing these self-modifying mental processes with Piaget's set of innate reflexes, however, we prefer to pair them with domain-specific skeletal structures. We do so for two reasons. First, there are now simply too many demonstrations of infant conceptual competence to accept Piaget's nonconceptual account of infancy. Second, reflexes do not provide an adequate foundation for learning: they cannot self-modify, but depend instead for their modification on other neural processes. For example, the vestibulo-ocular reflex (which keeps our eyes stable as our heads move) depends on separate neural machinery for calibration. When that separate machinery is removed, the reflex can no longer calibrate itself. By contrast, skeletal structures *do* provide an adequate foundation for subsequent learning because they can and do self-modify. If we accept the idea that from the start, our young apply their constructivist tendencies to the skeletal structures they are endowed with, then we have a way to account for how beginning learners come to share knowledge structures with their elders. Existing structures of mind help novices move onto learning paths that are domain-relevant rather than domain-irrelevant. Domain-relevant paths lead learners to inputs that are consistent with the structure of the domain to be learned; domain-irrelevant ones do not. More generally, the idea is that constraints function to enable and facilitate the acquisition and use of domain-relevant knowledge. Like other experientially modifiable biological structures (e.g., muscles), mental structures both require and contribute to epigenetic interactions with their environment (Dickinson & Dyer, 1996; Gottlieb, 1983; Lehrman, 1970).

Considerations of evolutionary, comparative, cultural, and Piagetian perspectives inform our assumptions that the mind has certain skeletal, innate domains of knowledge which actively contribute to its epigenesis of knowledge. Humans' ability to acquire novel knowledge in areas that

are not grounded on innate skeletal principles leads us to distinguish between *core* and *noncore* kinds of domains. We do not appeal to different learning mechanisms to explain the acquisition of each kind, however. For us, the key learning tool in both cases is structure-mapping, in which information from the environment is mapped with existing information in the mind. The mind's ever-present proclivity to find and map relevant data is enabled whenever there are already-present structures, no matter how skeletal they might be. Even skeletal structures can provide novices the wherewithal to find and map inputs that share their structure. Similarly, these structures can call on other mental learning tools, especially a frequency-computing tool, to collect information about the relative predictive validity of different cues as indicators of the appropriate interpretation or role assignment for an object/event. Skeletal structures give young minds a mental running start and serve as the engines of learning.

Where mental structures have to be acquired *de novo*—as is surely the case for topics such as chess, Newtonian mechanics, the theory of evolution, the stock market—learners have to acquire domain-relevant structure in addition to the content of the domain (Brown, 1990). Acquisition of noncore domains should therefore be difficult, since it is not easy to assemble truly new conceptual structures (e.g., Carey, 1991; Chi, 1992; Chi, Glaser, & Farr, 1988; Kuhn, 1970), and lengthy formal instruction is often required. Efforts to provide such instruction must recognize and overcome a crucial challenge: learners may assimilate inputs to existing conceptual structures even when those inputs are intended to force accommodation and conceptual change (Gelman, 1993, 1994; Slotta, Chi, & Joram, 1995).

To summarize, given that all theories of cognitive development and concept learning contain some notion of constraint, the preference for one class of theories over another should be guided by the degree to which facts about the mind in general and about cognitive development in particular are accommodated. We turn to marshaling the kinds of evidence that led us to a theory which embeds a learning-enabling notion of constraints.

KEY POSTULATES ABOUT LEARNING AND COGNITION

The last 25 years or so have provided a wealth of important data about cognition and cognitive development. Here we highlight seven postulates which attempt to capture theoretical advances and certain classes of findings that are

especially relevant to the consideration of different accounts of learning and concept acquisition. In the past, when appreciated at all, these postulates have commonly been considered individually or in small groups, rather than as a coherent and interconnected package. By bringing them all together here, we hope to build a solid theoretical and empirical standard against which subsequent discussions of constraints and their influence on learning can be measured. Where possible, we refer the reader to relevant chapters elsewhere in this *Handbook* for a greater degree of empirical detail. To anticipate the argument which we will develop later in this chapter, we believe that these postulates implicate a learning-enabling view of constraints within a rationalist-constructivist theory of mind.

Postulate 1: Inputs Can Be Divergently and Convergently Ambiguous, Requiring Interpretation

The would-be learner must cope with two fundamental challenges: perceptual and conceptual inputs are often either *divergently* or *convergently* ambiguous. Inputs that can be interpreted in more than one way are *divergently ambiguous* (the familiar sense of ambiguity). For example, many sentences are syntactically ambiguous, such as "They are baking apples," or "Jane likes Bill better than Will," or "Does he write on time?" Homophonic words are ambiguous by definition: the word "bat" can refer to an animate object, an inanimate object, or a verb. Notation systems provide other examples of divergent ambiguity. The notation $F(2,4)$ can stand for a fraction, a musical time signature, or even a statistical analysis-of-variance result. When used as a fraction, one can say that $\frac{2}{4}$ is equal to $\frac{100}{200}$ but not $\frac{1}{2}$. In contrast, the time signature $\frac{2}{4}$ can be rendered meaningfully into a $\frac{1}{2}$ rhythm, but there is no such thing as the time signature $\frac{100}{200}$. For a child learning both Spanish and English, the sound *si* can mean "yes" or "see" depending on whether it is meant to be interpreted as a word in Spanish or English.

Inputs can also be *convergently ambiguous* in their mapping to mental representations: that is, multiple perceptual inputs can converge on a common meaning. A single mental representation can have multiple perceptual instantiations. Bilingual language learners confront especially compelling examples of items which differ in their surface appearance but nevertheless have the same meaning. The French-English bilingual child has to master the fact that *un, deux, trois* and *one, two, three* share the same meaning, even though they are perceptually different. Turning to the mathematics of rational numbers again, the fractional

expressions $\frac{1}{2}$, $\frac{33}{66}$, $\frac{4}{8}$, and $.50$ are all equivalent and therefore share the same meaning, even though they look very different from one another. Similarly, the printed and written versions of individual letters vary extensively by case, font, and handwriting, even though they all share the same phonemic interpretation. When an adult and a child utter the same sentence, the much higher pitch range used by the child does not change the sentence's meaning. In general, much of the surface variability in what we perceive is discarded as irrelevant when we interpret its meaning.

Convergent and divergent ambiguities present a deep challenge to any learner, because they eliminate the possibility of a guaranteed, invariant, one-to-one mapping between a particular environmental input and a particular "correct" mental representation of that input that is to be learned. How do we learn what entities constitute the possible extensions and meaning of a concept? How do we learn particular mappings which are consistent from one occasion to the next, let alone consistently shared with other people with whom we might wish to communicate? These questions pose deep difficulties for any theory of conceptual development that seeks to build knowledge entirely from sensory and perceptual inputs.

Some readers might dispute our postulation of divergent and convergent ambiguity, denying the existence of these problems by arguing that the real world in which people live and act is contextually rich and full of structured information. There is no problem of stimulus impoverishment, so the argument continues, since people are able to use context to pick up and disambiguate relevant inputs (e.g., E. J. Gibson, 1984; J. J. Gibson, 1983; Greeno, Smith, & Moore, 1992; Lave & Wenger, 1991; Nelson, 1987, 1988; Rogoff & Chavayah, 1995; Rogoff & Lave, 1984). We believe that this response begs the question of how learners could accomplish this disambiguation. As can be seen in Figure 12.1, environments can present ambiguities about edges of objects, what is a "real" object as opposed to its reflection, and which occluded bits go together to form a whole object. The more novel and complex the physical environment, the more likely it is that objects and events will overlap. The environment by itself does not specify an interpretation, nor does it tell the mind what goes with what. We agree that there surely are cases where the structure of the environment affords or dictates particular solutions (picking up, pushing) and not others (twigs on trees do not afford standing-on). We particularly agree that relevant environmental inputs are often structured, a point to which we return below. However, we disagree that the world itself suffices to define what is relevant.



Figure 12.1 Where does what start (end) and what goes with what? Photograph by Adam H. Gallistel. Reproduced with permission.

If left to their own devices, novices are at risk for mixing together inputs that do not lead to a coherent perception or cognition, and for achieving perceptual organizations of the environment that differ radically from ones reached by others (Neisser, 1966). This is true even when the information source is social. For instance, contrary to our intuitions, the sound stream of spoken speech does not have gaps of silence in between words. Linguists have shown that this "segmentation problem" is a real challenge to language learners, because it is not obvious how to divide up the continuous speech stream into discrete words (let alone phonemes). Misinterpretations of well-known lyrics provide clear evidence of segmentation errors. Examples from both children and adults abound, including, "Gladly the

cross-eyed bear" (Gladly the cross I bear), "Our father Wishart in heaven" (Our Father, which art in heaven), and "One nation, underground, invisible" (One nation, under God, indivisible). In addition to demonstrating the problem of divergent ambiguity, these errors often reveal the action of top-down processes in children's active attempts to construct a coherent interpretation of what they hear. In the third example above, being underground and being invisible fit together well, giving the entire phrase a certain (unfortunately incorrect) coherence.

More generally, the notion of "context" as an ambiguous input's general surrounding is itself ambiguous. What counts as context for an ambiguous input can be adjacent inputs, or concepts, or points of view, or instructions, or knowledge of a language, or who else is present, or what one saw on the news, and so on indefinitely. Whether one "sees" a fossil or a stone depends on the prior knowledge and expertise one brings to the setting. What we "see" when faced with ambiguous inputs often depends on our mindset, on how our minds interpret a given surface input: we can "see" five individual animals or a cooperating group, an inclined plane or the affordance of something to slide on, a space between words or a fill-in-the-blank test item, a poor child or a dirty child, and so on.

The prevalence of divergent and convergent ambiguity points to an additional corollary: inputs which are incomplete can often still be constructed by the learner. What a child learns need not be fully grounded on what we as adults think are the relevant data "out there" in the world. Relevance is defined by the learner, not the teacher or the outside observer. For example, despite the seeming lack of "relevant" input, blind children learn that barriers interfere with visual perspective-taking in sighted individuals (Landau & Gleitman, 1985). Deaf children learn about complex aspects of the structure of American Sign Language (ASL) morphology which are not part of the input offered them by their signing parents (Johnson & Newport, 1991; Newport, 1990). Children create their own non-standard words that are nevertheless consistent with linguistic principles and rules (Clark, 1993; Marcus et al., 1992), as for example "mouses" (mice), "gursh" (ice cream, for one of our children), "furniture-animal" (statue), "undisappear" (to make appear again), "brealed" (broke), and "twenty-nine-ten" (thirty). Likewise, some children use idiosyncratic count lists that honor counting principles even though they differ from the conventional list of their language, for example, "1, 2, 3, 4, 5, 6, 7, H, I, J" (Gelman & Gallistel, 1978; Hartnett, 1991). A specific or complete set of inputs to be matched is not

required for learning to proceed. Learners can construct inputs that on their own are ambiguous or incomplete.

Postulate 2: The Mind Is an Active Learner, Not a Passive Receptacle

From the beginning, learners actively apply their motoric, sensory, perceptual, conceptual, and social-emotional capacities to their environments, immature as such capacities may be. Even babies have a ubiquitous tendency to use their limited repertoire of responses to explore and seek information from their world, including sucking, head-turning, eye-tracking, kicking, touching, crawling, standing, looking over or around barriers, etc. (e.g., Gibson, 1988; Haith, this volume; Piaget, 1952). Five- to 12-month-old infants not only like to look at pictures, they are willing to suck a pacifier at a particular rate in order to bring them into focus (Kalnins & Bruner, 1973). One-month-old infants will suck or turn their heads in order to hear speech sounds (Eimas, 1974; Mehler & Christophe, 1995), and will move their eyes around to explore objects (Banks & Salapatek, 1983; Haith, 1980). Infants also engage in actions which can have causal effects: they repeat an action to make interesting things last or reappear (Piaget, 1952). They kick in order to move a mobile attached to their foot by a string (e.g., Watson, 1987; Rovee-Collier & Bhatt, 1993), and they both start and stop attending to their caregivers who respond in kind (Stern, 1977; Tronick, 1989).

Toddlers manually explore the edges of an undulating surface such as a waterbed, and then get down and crawl on it (Gibson, Riccio, Stoffregen, Rosenberg, & Taormina, 1987). They spontaneously touch objects within one category in sequence before doing the same for those in another category, even when the categories are abstract ones like animate and inanimate (see Mandler, this volume). They point to things in their environment and ask "What's that?" sometimes so relentlessly as to tax the patience of their caretakers! As preschoolers, they persist at balancing blocks while exploring their physical characteristics (Karmiloff-Smith & Inhelder, 1974/1975). They spontaneously count things, including steps, cracks in sidewalks, puzzle pieces, raisins, train cars, and cows in the fields they pass on a trip (Gelman & Gallistel, 1978). They ask causal questions such as "why" and "how" (Koslowski & Winsor, 1981). It is hard to overemphasize the significance of these child-initiated perceptual and cognitive engagements of their environments. The following is an especially articulate defense by Eleanor Gibson of the active perceiver

(Gibson, 1988, p. 5), although see above for our differences with her views on the necessity of interpretation:

The old view of perception was that "input" from stimuli fell upon the retina, creating a meaningless image composed of unrelated elements. Static and momentary, this image had to be added to, interpreted in the light of past experiences, associated with other images, etc. Such a view of perception dies hard, but die it must. There is no shutter on the retina, no such thing as a static image. Furthermore, perceiving is active, a process of obtaining information about the world (J. J. Gibson, 1966). We don't simply see, we look.

Readers familiar with Piaget's writings can surely generate their own parallel passages to support the conclusion that neither child nor adult are passive receivers of conceptual inputs. We seek out and even create environments that nurture our active construction of knowledge (e.g., Piaget, 1971). As next illustrated, we also are very selective about what counts as relevant data for learning.

Postulate 3: Cognition and Learning Are Always Selective

The inbound flow of information-bearing signals from learners' sensory systems far exceeds their information processing capacities. We cannot perceive or learn about everything at once; by necessity, only a particular subset of the information with which we are continually bombarded can be perceived, learned, or thought about. Extensive research has been devoted to selectivity in vision, audition, attention, learning, and recall (e.g., Bregman, 1990; Broadbent, 1958; Garcia & Koelling, 1966; Trabasso & Bower, 1968; Treisman, 1982). We call attention here to just a few examples of this selectivity.

Selective Visual Attention

Neisser and his colleagues provide particularly compelling demonstrations of selective visual attention abilities in adults as well as 4-month-old infants (Bahrack, Walker, & Neisser, 1981; Neisser & Becklen, 1975). Subjects were shown two film events superimposed on each other, for example a handball game and a hand-clapping sequence. Care was taken to select movies that elicited comparable levels of attention when shown individually. Subjects were cued to attend to only one of the two events on the film: adults were told which one to attend to, and infants were played only one event's soundtrack (e.g., bouncing or clapping sounds).

Both studies demonstrated selective attention abilities. Adults readily followed instructions to monitor one of the superimposed events, and when asked to monitor both events at once, declared the task difficult or impossible. Selective attention in the infants was demonstrated by the method of habituation, which takes advantage of infants' preference for novelty. Having repeatedly seen the superimposed films while hearing only one of the two soundtracks, infants were tested by being shown the two movies silently and separated side-by-side. In this test situation, the infants looked reliably more at the movie that was silent during the initial phase, presumably because they had not been attending to it previously and it was therefore novel to them.

These selection effects were among the results used to support Neisser's (1966) view of perception in both infants and adults, in which schemata are formed and actively tuned based on initial uptakes of relevant, attended-to information. Information about irrelevant events is not noticed, for the simple reason that no schemata for processing such events are generated.

Selective Learning in Spatial Localization

Studies of how rats (Cheng, 1986) and toddlers (Hermer & Spelke, 1994; Spelke & Hermer, 1995) learn object locations in space add strength to the proposal that schemata and conceptual models are responsible for selective "screening-out" effects in learning. Both rats and toddlers use certain types of information to locate themselves in space, but not others. For example, Cheng made concerted efforts to encourage rats to use various types of landmark information when orienting themselves in space: odor, texture, and surface information presented at the same place as food. None of these three types of information were learned, despite their high salience to humans: rats preferred instead to use the overall shape of the experimental environment. This reliance on overall shape was demonstrated even in situations where shape was a particularly uninformative cue, for instance in a rectangular four-arm radial maze. In one such study, after learning a particular food location (e.g., a well-marked corner) and then being disoriented and placed at the center of the rectangular box, the animals chose randomly between the correct corner and the one 180° opposite. They did this despite the perfect association between the food and nearby background cues. Had they used these cues, they should have made few errors.

Spelke and Hermer (1995) report the same response pattern in studies with toddlers. Toddlers first watched as a toy was hidden in one corner of a rectangular room. They

were then blindfolded and turned several times before being allowed to search for the hidden toy. As with Cheng's rats, seemingly salient landmark information (a blue cloth covering one end of the room) was ignored in favor of ambiguous geometric information about the overall shape of the environment (e.g., "it's in the corner"). "Obvious" pairings were not learned (or at least, not sufficiently to guide successful searching behavior). The children looked for the toys equally often at geometrically equivalent places (diagonally opposite corners of the rectangular room).

Selective Learning and Domain-Specific Cue Salience

Studies of the learning conditions under which young children do or do not use color information provide another major example of selectivity. We know that infants perceive colors categorically, that is, like adults they divide the continuous color spectrum into discrete categories such as red and blue (Bornstein, 1985). Further, there are ample opportunities for young children to encounter differently colored blocks, letters, toys, stacking cups, pieces of paper, stickers, and so on. Thus, it would seem that the conditions for acquiring color-relevant concepts and knowledge are met at an early age. Nevertheless, preschoolers have substantial difficulty learning the names for different colors, and they do not use salient color differences in many concept learning tasks (for reviews, see Bornstein, 1985; Landau & Gleitman, 1985; Macario, 1991). By contrast, at the same age, these children do learn quite a few number names (Shatz & Backscheider, 1991).

Hirschfeld's (1996) studies of how preschoolers assign individuals to different categories provide another example in which color is not a significant variable. Children did not base their classifications on skin color: again, the suggestion is that this inattention follows from the children's own ideas about what is relevant and irrelevant. Macario's (1991) studies of 2- to 5-year-olds' use of color are especially interesting in this regard. When children were presented with novel objects and thought they were learning about new foods, they used the color information provided with ease. When the same novel objects were thought to be exemplars of new toys, however, color was treated as irrelevant.

These demonstrations of the variable salience of color information share a common theme. The fact that a perceptual attribute is ubiquitously present in the environment does not guarantee its psychological relevance, nor its ready uptake for perceptual processing and concept learning. Relevance at the psychological level is very much a matter of structure, a point to which we will return below

in Postulate 7. Selectivity in learning and cognition is a necessary byproduct of sensory overload. How do we define what is relevant, and choose the particular information subset we do? Selectivity can be accomplished either randomly or nonrandomly, and random selectivity is inconsistent with what we know about commonalities in learning paths and outcomes across individuals and cultures. The remaining alternative is nonrandom selectivity, and our preferred explanation emphasizes the role of enabling constraints in defining relevance within core domains.

Postulate 4: Animals, Infants, and Young Children Are Not Perception-Bound Learners

We draw attention to an important feature of our discussions of selective attention and structured inputs. Many of our examples show infants and very young children selectively attending to and using inputs that are more abstract than concrete (see also Postulate 7). We surely have not exhausted the list of such examples. For example, it is now known that young children appeal to invisible entities in explaining their belief that something is contaminated (Au, Sidle, & Rollins, 1993); they invoke internal or invisible causal forces to explain why objects move and stop (S. Gelman & Gottfried, 1996; Massey & Gelman, 1988; Williams & Gelman, 1995); they reason about internal body parts and their functions (Carey, 1985) or the contrast between the insides and outsides of unfamiliar animals (Gelman, 1990; Massey, 1989); and they can readily pretend that the same empty cup is first a full cup and then an empty cup (Leslie, 1994a). In this section we extend our review of conceptual abilities to encompass animal cognition. More and more students of animal behavior are demonstrating that animals have nonverbal conceptual abilities and are not limited to instinctual sensing, perceiving, and reacting. As a result of ethological and experimental research with animals over the past fifty years, we now have solid evidence that many animal species have evolved complex representational abilities related to solving particular adaptive problems (see reviews in Gallistel, 1990, 1995; Gallistel et al., 1991).

The demonstrated abilities of animals and infants to use abstract information and acquire abstract concepts are inconsistent with the long-standing view that they are pre-conceptual, "perception-bound," and limited to concrete reasoning or stimulus-response pairings. Gallistel (1990) provides a review of the findings showing that all kinds of animals orient themselves in space with reference to a cognitive map, that is, a representation of the relative positions of points, lines, and surfaces in their surroundings. He also

reviews the extensive studies demonstrating the surprising abilities of rats and pigeons to represent numerosity and temporal intervals, as well as to perform common arithmetic operations upon these quantities such as computing sums and ratios. Below, we introduce two highlights from these remarkable findings: evidence for numerical abilities (see also Postulate 7 for a discussion of numerical abilities in infants), and evidence for cognitive-mapping abilities.

Since several chapters in this Volume provide extensive coverage of the young child's considerable reasoning and representational abilities (DeLoache, Miller, & Pierroutsakos, Ch. 16; Mandler, Ch. 6; Wellman & S. Gelman, Ch. 11, this Volume), we will not dwell on them further (see also Gelman & Baillargeon, 1983). Instead, we describe a compelling example of abstract reasoning and concept acquisition by infants: mechanical/biomechanical reasoning about inanimate and animate objects.

Numerical Abilities in Animals

Platt and Johnson (1971) provide an excellent demonstration of rats' numerical abilities. The animals in their experiment were required to press a lever a certain number of times ($N = 4-24$) to arm a feeder that operated silently. After the N presses on the lever, the rat could poke its head into a feeding alcove and obtain food. If the rat failed to press the lever the requisite number of times before poking his head into the alcove, the counter was reset to zero. The observed median number of presses corresponding to the required value of N , showing that rats can count up to at least 24 (nonverbally, of course). Platt and Johnson's study is but one of many examples of animals counting beyond very small numbers (e.g., Capaldi & Miller, 1988; Meck & Church, 1984), indeed for N as large as 50 (Rilling, 1967). These data, together with demonstrations of the chimpanzee's and parrot's abilities to use numerical symbols in arithmetically correct ways (Boysen & Bernston, 1989; Pepperberg, 1987), have encouraged the view that animals can and do have abstract cognitions, despite the fact that they do not possess a rich linguistic capacity.

Cognitive Maps in Animals: More on Spatial Abilities

In our discussion of selectivity under Postulate 3 above, we introduced evidence that rats can form cognitive maps of their environment. A compelling earlier demonstration of this ability comes from Morris (1981). Rats were dropped into a pool of opaque water, in which they had to swim until they found a barely submerged brick on which to perch. On subsequent trials, they were able to swim directly to the perch, whether or not it had a flag placed on it

to identify its location. This kind of orientation would not be possible if the rat could not place the position of its goal within the framework of a cognitive map of the macroscopic shape of its environment.

An additional example of cognitive mapping abilities in the absence of immediate perceptual information comes from research with goby fish (Aronson, 1951). These fish live in tide pools, and can become stranded by receding tides. In order to escape to the open sea, they accurately leap out of the water from pool to pool, despite being unable to see the destination of their jumps in advance. Through experimental manipulations with artificial tide pools, Aronson demonstrated that jumping accuracy depended on experience during a rising tide which allowed the fish to learn a cognitive map of the spatial arrangement of the pools.

Conceptual Abilities in Infants: "Baby Object Mechanics and Biomechanics"

Work on infants' knowledge of objects and how they move adds strength to the idea that concept learning can go from the abstract to the particular. Very young infants have been shown to respond in ways which are consistent with highly abstract principles about the nature of moveable physical objects, distinguishing between those that are animate and those that are inanimate (Mandler & Wellman, Ch. 6, this Volume; Sperber, Premack, & Premack, 1996). In general, these demonstrations have found that infants look longer at "impossible events" (in which objects appear to violate physical laws) than they do at "possible events." For example, 5-month-old infants respond in ways consistent with the beliefs that one solid object, a rotating screen, cannot pass through another solid object, a block hidden behind the screen (Baillargeon, Spelke, & Wasserman, 1985) and that an inanimate object cannot propel itself (Leslie, 1995).

Spelke (1991) provides similar evidence with 3- and 4-month-old infants, in a habituation study for which both the possible and impossible events were novel. To start each habituation trial, a ball was held above a screen and then dropped behind it, after which the screen was removed to reveal the ball resting on the surface of the table. This event sequence was repeated until infants habituated, that is until they looked less than half as long during a trial than they once did. During the posthabituation trials, the ball was again dropped, but this time the screen was removed after the drop to reveal the object resting either on top of or underneath a novel shelf that had been placed surreptitiously into the display. For the latter case, the ball ended up in a familiar position, on top of the table. But to get there, it

would have had to pass through the shelf that sat between the dropping point and the table. Therefore, although both of the posthabituation events were novel, only the latter event (ball-on-table) was impossible. Once again, infants looked longer at the impossible event. In a control set of conditions, infants saw the same final displays in the posthabituation phase but the balls were not dropped. Therefore, neither outcome was impossible. Now infants preferred to look at the ball that was resting in a novel place, that is, on the shelf as opposed to the table. A series of follow-up studies by Baillargeon and colleagues have confirmed this finding, extended it to still younger infants (Needham & Baillargeon, 1993), and also investigated infants' understanding of physical phenomena including support, collision, and unveiling (Baillargeon, 1995; Baillargeon, Kotovsky, & Needham, 1995).

These findings are joined by a series of studies demonstrating infants' ability to distinguish between novel examples of mechanical and biomechanical motion (Bertenthal, 1993). Infants distinguish between objects and events which are animate or inanimate on the basis of differences in the conditions that cause movement. Spelke, Phillips, and Woodward (1995) provide especially compelling evidence here since they controlled for differences in size and surface characteristics between animates and inanimates. The authors contrasted infants' reactions to both animate and inanimate pairs of videotaped displays. Inanimate pairs of stimuli were two 5- and 6-foot-tall objects that had distinctive novel shapes and contrasting bright colors and patterns. The animate pairs were two people. With each pair of objects, infants watched two events: (a) the objects moved toward each other, touched each other, and then changed direction; and (b) the objects moved toward each other, stopped briefly before touching each other, and then changed direction. During test trials, infants looked reliably longer at the no-contact inanimate event; they showed no such preference in the animate event trials. This higher looking time is usually an indication of either surprise or a violated expectation. The authors' conclusion is that 7-month-olds know that a causal event between two inanimate objects requires contact, but the same is not true of two animate objects. In other words, animates are capable of action at a distance, but inanimates are not. This pattern of results is exactly what one should observe if the innards and external-agent causal principles proposed by Gelman (1990) aid infants in interpreting the motion paths of animate and inanimate objects.

Converging evidence that infants interpret the same motion path differently, depending on whether the moving

object is animate in kind or not, can be found in Golinkoff and Harding (1980), Poulin-Dubois and Shultz (1989), and Premack and Premack (1995). This line of work is joined by studies showing an early ability to distinguish animate from inanimate objects and to distinguish the causal conditions for their differing kinds of motion and transformations. For instance, Mandler and McDonough (1996; this volume) have demonstrated that 7-month-old infants can categorize replicas of animals and nonanimals, as revealed by their reliable tendency to touch items within one abstract category before switching to explore those in another.

There are several reports of young children applying a "no action at a distance" principle of physical causality for inanimate objects but not animate ones (e.g., Bullock & Gelman, 1979; Bullock, Gelman, & Baillargeon, 1982; Koslowski, Spilton, & Snipper, 1981). When the 3- and 4-year-old children in the Bullock and Gelman studies were shown an event that had a gap between the cause and outcome (cf., Leslie & Keeble, 1987), they either inferred a mechanism ("When I wasn't looking, the ball slid over"), talked about magic, or made it clear that something was not quite right ("What? How did that happen? It's a trick, right?"). Similarly, we find that young children produce predominantly external causal attributions for the motion of familiar inanimate objects but not the motion of familiar animate objects (S. Gelman & Gottfried, 1996; Massey & Gelman, 1988; Williams & Gelman, 1995). For animate objects, children offered attributions that were related to the object class itself: these included the object's capacity for various self-propelling actions, its limbs, and some combination of blood, bones, and food. When this pattern is compared with their answers to questions about the insides of animates (Gelman, 1990), we are reminded of René Descartes' suggestion that animate action resembles the motions created by the hydraulic system controlling the statues in the Gardens of Versailles, and of Claude Bernard's description of people as nothing but a bag of bones filled with and surrounded by liquids and foods. Hence, we are much inclined to limit ourselves to an account of the early conceptual understanding of animates in terms of a theory of internal mechanistic action (Williams & Gelman, 1995), rather than an early theory of biology (cf., Au & Romo, 1996; Carey, 1995).

In sum, how can we account for young novices' selective and successful use of abstract data before they have induced the necessary abstractions in the first place? How do 5- to 7-month-old infants know to perceive simple collision events in terms of the causal roles played by the striking

object (agent) and the struck object (recipient), rather than solely in terms of the spatiotemporal relations between the objects (Leslie & Keeble, 1987; Leslie, 1995)? The existence of so many abstract cognitive abilities in infants, young children, and animals requires us to answer this fundamentally important question. One way to do this for the findings on number, causality, and the animate-inanimate distinction is to propose that domain-specific skeletal principles enable construction of a model of the world. If these model-building principles define the structure of a domain, and if the mind actively uses whatever structure it has as a basis of induction, then inputs that are related in complex ways to the inductions drawn from them will be favored.

Postulate 5: Learning Is Not a Single "Process" Since There Are Different Domain-Specific Kinds of Learning

We now know of an ever-increasing number of cases in which the rate of learning varies as a function of what is being learned and when. Some learning proceeds smoothly and at a relatively rapid pace. The course and conditions of such early learning in core domains can contrast strongly with the slow and uncertain learning in noncore domains of knowledge—even when we think a core and noncore domain are closely related. Studies of complex learnings in animals have been central in the development of domain-specific accounts of animal learning (e.g., Gould & Marler, 1987; Shettleworth, 1993) that were fueled by demonstrations of species-specific learnings about illness (Garcia & Koelling, 1966; Rozin, 1976) and danger signals (Bolles, 1970). These types of learning include bird song, categories of predators, the time of day at which events occur, the rates of event occurrence, and navigational information such as the solar ephemeris and the center of rotation of the night sky. Many of the abilities and behaviors exhibited by animals cannot be explained by traditional behaviorist or associationist learning models (Gallistel, 1990, 1995; Gallistel et al., 1991; Rozin & Schull, 1988).

Theoretical developments within evolutionary and animal learning theory have encouraged us and other students of conceptual development to adopt a key idea from evolutionary theory: learning is the product of behavioral mechanisms with elaborated internal structures that have evolved to guide domain- and species-relevant learning (see Carey & Gelman, 1991; Hirschfeld & Gelman, 1994). Additional examples of such learning in humans (see also above) include foods which are safe to eat (Siegal, 1995), the faces of others in a newborn infant's world (Johnson &

Morton, 1991), the syntax of language (Pinker, 1991), and probably implicit rules of conversation (Grice, 1975).

Within this theoretical frame of reference, concept learning is not a homogeneous or content-neutral process. Rather, it benefits from multiple mental structures that support attention to and learning about domain-relevant, structured data sets (Cosmides & Tooby, 1994a; Fodor, 1972, 1975; Gallistel, 1995; Gardner, 1992; S. Gelman, 1991, 1993; Hirschfeld & S. Gelman, 1994; Leslie, 1995; Medin, Goldstone, & Gentner, 1993). This view of learning differs in fundamental ways from theories grounded on general processes, for example, Skinner (1938, 1950), Hull (1952), Thorndike (1913), Watson (1913), Rumelhart and McClelland (1986), Simon (1972), and Siegler (1991). Information processes embodied in abstract domain-specific structures privilege attention to domain-relevant, structured data, an especially important idea which we develop at length below. This means a critical condition is met for learning; the learner's active structure-mapping tendencies (see below) can be put to work in the service of early concept development and learning, no matter what specific learning theory we embrace.

To say that knowledge acquisition in core domains benefits from innate skeletal structures is not to say that there is no need for learning. The knowledge in question will not appear full-blown immediately when the individual encounters an example of relevant data. We do not hold such a position, nor do most other theorists working within a domain-specific framework. Our goal is to achieve an account of learning that better handles the evidence. Ours is not a theory that falls within the empiricist class of theories about cognitive development and concept acquisition; nevertheless, it is a theory of learning. Variability is a fundamental fact about cognitive development.

Postulate 6: Epigenesis Leads to Variability and Requires Supporting Environments

With rare exception, any genetic developmental program carries with it extensive requirements for interactions with environments that can nurture, support, and channel the differentiation of adult structure. In the absence of those environments, the program will almost certainly fail. The same is surely true for skeletal mental structures: the existence of a primordial input-structuring mechanism does not guarantee that related knowledge will spring forth full blown the moment the individual encounters a single example of the requisite environment. Without opportunities to

interact with, learn about, and construct domain-relevant inputs, as well as to practice components of relevant action plans, the contributions of skeletal structures will remain unrealized, or will lead to atypical developments. It follows that learners must encounter opportunities to interact with and assimilate relevant supporting environments (cf., Scarr, 1993). It also follows that variability is a characteristic of any learning, whether about core domains that benefit from skeletal structures or noncore domains that do not.

Marler's work on birdsong learning provides an especially compelling example of epigenetic variability from the animal literature. His work on the course of songbird learning in the male white-crowned sparrow provides us with an important case of animal learning in which skeletal structures guide learning but also lead to variability. These findings in combination with ones from several developmental laboratories illustrate the point of this postulate: be the learning in a core or noncore domain, variability is to be expected, and is entirely consistent with a theory that invokes *enabling* constraints on learning.

Songbird learning mechanisms are specific adaptations, designed to operate in environmentally specific contexts to ensure the uptake of adaptive information. Innately-determined representations guide selection of what is to be learned, and how that subset is both extracted from diverse input and selectively remembered in constructing a song. Of major interest here is the fact that song sparrows who grow up in different places end up singing different songs, namely, songs with the characteristics of the dialect they are exposed to plus individual signature elements. White-crowned sparrows, like many songbirds, have a sensitive period somewhere during the first year of their life. If they do not get to hear the song of an adult sparrow during this time, the song they sing will be decidedly odd when they grow up. The sensitive period has the effect of "tuning" the bird's template for a particular class of song. But exposure during an early period of life does not suffice to guarantee that the young bird will grow up to sing the represented song, or that his adult song will be normal. He has to go through periods of subsong learning and crystallization, again during the first year of life.

The subsong learning period is especially interesting for us, because only at this time does the young bird start to sing (after the sensitive period for exposure). The initial efforts look very much like trial-and-error, and in some ways are like a human infant's babbling; the song sounds not at all like the adult target model. The data gathered by Marler and colleagues compel the conclusion that the bird

is working at converging on a particular output plan, one that generates a song that is consistent with the mentally represented song template. Without any more input than what the young bird creates himself at this point, the bird gradually moves toward singing a song that contains more and more notes and phrasal units of the adult song (Marler, 1991; Nelson & Marler, 1993). Yet because of the idiosyncratic nature of this learning process, no two adult songs are identical: there is variability in the outcome.

It is hard to resist comparing the young white-crowned sparrow's learning path with examples of children's learning. One case that seems especially pertinent comes from beginning language learners' efforts to master the count list of their language. The evidence is that children detect the relevant list of count words at a very young age (Gelman, 1990; Shatz & Backscheider, 1991; Wynn, 1992). Yet only much later will they reliably produce even the first ten words in the list and use this verbal knowledge without error (Fuson, 1988; Hartnett, 1991; Miller, Smith, Zhu, & Zhang, 1995; Wynn, 1992). Karmiloff-Smith and Inhelder's (1974–1975) microgenetic analyses of children's block-balancing show similar developmental patterns, from seeming trial-and-error-like efforts to systematic, rule-governed actions. We are cautious about labeling these early counting and block-balancing efforts as examples of trial and error, however, for even this variable performance is domain-relevant and far from random. Siegler and Crowley's (1994) experiments on how 5-year-olds achieve arithmetic solutions by converging on the use of a novel strategy, and how 10-year-olds learn to succeed at tic-tac-toe, are particularly relevant in this regard. They provide elegant demonstrations of the importance of paying special attention to learners' goals and abilities to attempt nonrandom solutions en route to a successful solution.

Like Siegler and colleagues (Siegler, 1991; Siegler & Crowley, 1991), Kuhn's (1995) review of a variety of microgenetic studies examining change processes in learning concludes that variability is a fact about learning, no matter what one's theory of learning might be. Systematic within- and across-condition variability in the extent to which performance conforms to abstract principles is consistent with traditional learning and developmental theories in which unprincipled "habits" are acquired prior to the induction of principles. But contrary to widespread assumption, such variability is equally consistent with and predicted by rationalist accounts of cognitive development. It therefore behooves us to consider potential sources of variability more carefully, and examine how different

theories account for these. Variability could conceivably result from random noise in learning mechanisms, different learning solutions for the same underlying structure, different task demands, cultural differences in the interpretation of test settings, differences in the development of planning abilities, or a lack of achieved competence. We will return to the issue of variability below and the possibility that there may be different but systematic sources of variability that are due to conceptual learning goals, cultures, physical environments, settings, tasks, etc. (Cole, 1995; Donaldson, 1978; Gelman & Greeno, 1989; Gelman, 1993; Rogoff & Chavayah, 1995; Siegal, 1991).

Postulate 7: Environments and Minds Share Universal Structures, Not Universal Surface Characteristics

In our final postulate, we bring together elements from several of the foregoing postulates. In core domains, learners apply enabling structures in order to map structured information from the environment onto skeletal mental structures. We emphasize two aspects of the phrase "structured information from the environment." The first is that such information can be abstract rather than directly perceivable or concrete; nevertheless, the active mind is capable of selectively seeking out and learning the relevant abstract information. The second is that skeletal mental structures are attuned to information in the environment at the level of structural universals, not the level of surface characteristics. This means that enabling mental structures can be universal and yet still produce variability across individuals and cultures. We will develop these arguments by going into deeper empirical detail than in our discussion of the previous postulates. Our focus will be on research in a highly abstract area which we believe to be a core domain: numerical reasoning.

More and more we are gaining evidence that children all over the world acquire knowledge in core domains without explicit instruction, just as they do with their native language. This learning takes place automatically, given of course that they have access to the everyday cultural, social, and physical environments of their community (Carey & Gelman, 1991; Hirschfeld & S. Gelman, 1994; Hirschfeld, 1996; Keil, 1981; Mandler, this volume; Leslie, 1994b, 1995; Premack, 1990; Sperber, Premack, & Premack, 1996). In the core domain of numerical reasoning, all children develop a principled understanding of verbal counting and use it to add and subtract, whether or

not they receive formal schooling (for reviews, see Gelman & Meck, 1992; Miller et al., 1995; Resnick, 1989; Sophian 1994). Children all over the world also quickly learn about the differences between animate and inanimate objects (Atran, 1994; Carey, 1995; Inagaki & Hatano, 1987; Jeyifous, 1986; Simons & Keil, 1995; Wellman & S. Gelman, Ch. 11, this Volume).

In both of these cases there are differences about the details learned, for instance in the particular count list and the particular kinds of animals and nonanimals. It could not be otherwise, given that different cultures use different languages and live in different parts of the world. These kinds of findings challenge any account of development that builds knowledge from bits of sensory inputs, but they fit well with an account that sees young learners as using abstract domain-specific structures. As long as the encountered exemplars are consistent with the structure that outlines the domain, they can serve as domain-relevant food for thought. Number-specific principles can identify and assimilate the examples of count lists in one's native language. Similarly, causal principles that distinguish between animates and inanimates can encourage attention to the kinds of information that will support learning about which particular moveable objects are self-moveable and which need an agent to move them (Gelman, Durgin, & Kaufman, 1995; Premack & Premack, 1995).

From this perspective we should and do find that young children are able to use information about objects other than simply their surface perceptual features (see Postulate 4). For example, Massey and Gelman (1988) found that 3- and 4-year-old children do not always use overall surface similarities when judging which of a series of novel items can move both up and down a hill. Moreover, when justifying their yes-no answers, they often either denied visible information when it conflicted with their conceptually-based classifications, or invented information that was not perceptually available. Two examples were the following: "it does not have feet" (for a statue whose "feet" were clearly visible), and "it uses the feet" (for an animal whose feet were not visible). Children ignored obvious but superficial similarities between a statue and the living thing whose form it reproduced, treating the statue more like the wheeled objects in the study than the animate objects. Simons and Keil (1995), S. Gelman and Gottfried (1996), and Kremer and S. Gelman (1991) all report a similar preference among young children for abstract as opposed to surface information when sorting or reasoning about novel examples of animate and inanimate objects.

The conclusion that early learners use structured data is reinforced by the generality of the demonstrations of infants' conceptual abilities. These results have been obtained across a wide range of samples, including infants from the culturally diverse and low-income areas of downtown New York City, Philadelphia, Washington, and London, as well as more homogeneous areas such as Champaign and Ithaca (in the United States), Geneva, and Japan (see chapters in this volume by Banks & Kellman, and Haith). An especially interesting set of studies are those on number.

Young Children's Alleged Numerical Incompetence

Not so long ago, most developmental psychologists and educators concurred that children's ability to achieve numerical concepts was relatively late to develop. Preverbal infants and nonverbal animals certainly lacked numerical abilities, given that they were restricted to processing data at a sensorimotor or perceptual level. Young children lacking concrete operational structures were therefore also limited to the perceptual level. For example, they failed in Piaget's (1952) conservation of number experiment because they could not help but focus on perceptual attributes as opposed to the shared one-to-one correspondence of the two arrays. Although the claim that "primitive" people and young children lack abstract numerical abilities is less prevalent than it once was (see Crump, 1990), the theme continues (see, e.g., Andrews, 1977; Bidell & Fischer, 1992; Hallpike, 1979; Hurford, 1987; Ifrah, 1985; McCleish, 1992, for a discussion of this theme within fields of education). The following excerpt is but one of many examples that could be cited:

Even human beings, for whom the abstract activity of counting seems the most natural thing in the world, find it incredibly difficult to learn. One of the discoveries of the 20th-century scientists (made independently by Montessori, Piaget, and Vygotsky) was that adults forget how gradual and time-consuming the process of learning to think in the abstract is. (McCleish, 1992, p. 8)

Subitizing versus Counting: A Perception-Only Account of Number After All?

Several studies of infants' use of numerical information have revealed a set size limitation, in which infants reason accurately with displays of no more than 3 or 4 items. These results encourage many to continue to argue that infants use non-numerical representations rather than abstract numerical representations. The favored mechanism

is called "perceptual subitizing," a process that is assumed to allow subjects to make discriminations between small set sizes without any implicit or explicit understanding of numerical principles (e.g., Cooper, 1984; Cooper, Campbell, & Blevins, 1983).

It is true that ultimate mastery of the count list can take a very long time (Hartnett, 1991; Miller et al., 1995). However, this alone is insufficient reason to reject the proposal that the learning process is guided by implicit counting and arithmetic reasoning principles. It is important to distinguish between the ease of finding a learning path at all and the final mastery of the knowledge on that path (see Postulate 6 for a discussion of factors contributing to variability in count list learning). Similarly, although it is true that infants' ability to discriminate numerosities appears limited to small set sizes, it is far from clear that a perception-only model can accommodate the data.

The problem for perception-only models is that infants can process numerical displays with an extremely diverse range of object and event types, enough so as to constitute an abstract category. Starkey, Spelke, and Gelman (1990) showed photographs of common household items to 6- to 8-month-old infants, including a comb, pipe, lemon, scissors, and corkscrew. These items varied in color, shape, size, and texture. Each photograph showed two or three different items, and the spatial arrangement of the items was unique from trial to trial. Thus, the only common characteristic of the 3-item and 2-item displays was their numerical value. To start, half the infants were habituated to 3-item displays, half to 2-item displays. Despite all the variations in the objects' surface characteristics, infants were able to attend to number (set size): recovery from habituation occurred when infants saw a display with the other set size (e.g., when seeing a 3-item display after having been habituated to a series of 2-item displays). Wynn (1995) has demonstrated that infants can attend to numerical information not only with concrete objects but with dynamic events. Events are particularly difficult to attend to within an associationist learning theory because they occur across time, rather than consisting of co-occurring sensations that can be associated. In Wynn's studies, infants were able to keep track of the number of times a puppet jumped up and down, for instance dishabituating to a two-jump event after having been habituated to a series of three-jump events. Finally, an especially interesting demonstration of infants' ability to notice abstract number information in the environment is presented by Canfield and Smith (1996). These authors report that 5-month-old infants' anticipatory eye-movements reflect an

ability to keep track of the number of pictures presented in one location and use this running count to predict whether the next picture will be at that same location or another one.

Infants in these various studies could have focused on perceptual attributes of the items such as shape, motion, textural complexity, and so on, but they did not. To date, there is no known simple perceptual mechanism that can do this, that is, attend to number across an ever-increasing set size while ignoring item kind, color, size, shape, location, event type, degree of stimulus overlap, a wide range of visual angles, modality, the difference between simultaneous and sequential presentations, and so on. More and more it becomes difficult to continue to maintain that infants are restricted to some simple, low-level sensory or perceptual mechanism.

Because infants can process numerical information across sensory modalities (Starkey et al., 1990) and across time (Wynn, 1995), a mechanism that selectively attends to number must deal with the things being counted at a level far removed from their basic sensory or configurational properties. Two kinds of current explanations meet this requirement. One is that the infants use something akin to a set theoretic definition of "thing," where "thing" is a member of the class of all classes of objects and events. Numerical abilities are then attributed to the use of a one-to-one correspondence rule (see also Starkey & Cooper, 1995). The other is that infants use a less rarefied notion of "thing," the one embedded in the nonverbal counting principles of one-to-one, stable-ordering, and cardinality in the counting plus arithmetic reasoning model proposed by Gelman and Gallistel (1978) and Gallistel and Gelman (1992). If counts are governed by an implicit understanding of the how-to-count principles, then it follows that counting will be indifferent to item type (Gelman & Greeno, 1989). All that matters is that the collection of items in a to-be-counted set are perceptually or conceptually separable from one another. From this point of view, a "thing" is as basic a concept as one can specify (Gallistel & Gelman, 1990), reflecting the differentiation of figure/event from background, which is widely recognized to be one of the most fundamental operations in perception and cognition (Kellman & Banks, Ch. 3, this Volume; Spelke, 1990). In this account, infants need not even notice the color or kind of things they encounter when they are engaged in numerical processing (Xu & Carey, 1996). The numerical abilities follow from application of a structure that relates the counting principles to the effects of addition and subtraction.

One further line of argument is in order before leaving the issue of subitizing versus counting. Conclusions that infants use a perception-only "subitizing" mechanism are based in large part on a methodological decision to score infants' numerical discriminations as correct (exact) or not, rather than looking at the variance in errors that are produced for larger set sizes. This method is unable to determine whether errors are due to an inability to use a preverbal counting process, or instead to increases in inherent variability. Thus, the claim that infants cannot deal at all with larger numbers is problematic. To distinguish between these alternatives, it will be necessary to find ways to obtain infants' estimates of variability as a function of set size. This has not yet been done, but corresponding studies have already been done with animals that reveal an ability to use set sizes much greater than 3 or 4 (see Postulate 4 for more on animals' numerical abilities). Since the infant data that do exist in other areas of numerical reasoning map well onto the animal data, Gallistel and Gelman (1992) have proposed that infants use a mechanism like the Meck and Church (1984) animal counting model.

In the Platt and Johnson (1971) studies of rat counting described above, the greater the particular numerosity to be represented, the more likely it was that the animals confused it with adjacent or nearby values of N . The fact that the variance of the distribution for each N increases with N in all of the animal counting experiments has implications for how studies with infants are interpreted. We attribute this increase in variance to two systematic sources (Gelman, 1993). The first is generated by the counting process itself: increasing set sizes involve greater information processing demands and therefore an increasing chance of errors (Gelman & Gallistel, 1978; Gelman & Tucker, 1975). The second source is the scalar variability in the mental magnitudes with which animal and human brains represent scalar variables like number and the duration of temporal intervals (Gallistel, 1992; Gibbon, Church, & Meck, 1984). The coefficient of variation in these quantity-symbolizing magnitudes is constant, that is, the standard deviation of the accuracy with which a mental magnitude may be read increases in proportion to the magnitude. This coefficient of variation may be greater in young children than in adults. Thus, the variability in the representation of a given numerosity increases with increasing numerosity, both because longer counts have more counting errors and because the mental magnitudes by which numerosities are represented have intrinsic scalar variability. This increase in the variability or uncertainty means that something like Weber's law applies to the discrimination of remembered

numerosities: for a fixed difference between two numerosities, the bigger they are, the proportionally more difficult they are to discriminate.

In sum, infants' errors at set sizes where $N \leq 3$ may not implicate subitizing or a lack of numerical reasoning ability, but instead may indicate the presence of one common preverbal counting mechanism similar to the one used by animals. Indeed, Balakrishnan and Ashby's (1992) analyses of a huge number of trials generated by adult subjects over a two-week period show that there is good reason to assume that the same holds for adults. These authors conclude that there is but one underlying process controlling adult judgments of set sizes that are both smaller and greater than three or four.

Numerical Reasoning Beyond Counting

A key to the question of whether infants possess nonverbal numerical principles is whether they can reason about number with set sizes, even if no greater than three or four (Gelman, 1972). More and more it appears that infants are capable of reasoning about the effects of the arithmetic operations of addition and subtraction. For example, Wynn's (1992) 5-month-old infants responded appropriately when they saw addition and subtraction. The infants first had repeated experience with two objects; then a screen covered the objects and the infants watched as an object was added to or removed from the hidden display. The screen was then removed, revealing one or two items; in each case, infants looked longer at the numerically incorrect display. Data like these implicate a process that relates the effect of adding or removing items to a numerical representation of the initial display (Wynn, 1992, 1995).

Evidence of infants' ability to use the numerical ordering relations ($<$, $>$) strengthens the conclusion that they can reason numerically (for more on this point, see Gelman & Gallistel, 1978, Chapter 11). Although we know of no such data for infants as young as Wynn's, they do exist for older infants. Cooper (1984) showed that 12-month-olds can learn to compare successive visual displays on the basis of whether they have an equal or unequal number of items. Sophian and Adams (1987) took advantage of 14-month-old infants' preference for the larger of two sets. To start, they showed infants two displays with the same number of items in each. Then the displays were both screened and infants watched as an item was either added to, or subtracted from, one of the displays. When allowed to reach out, they had a reliable preference for the screen hiding the effect of subtraction, demonstrating their ability to relate the equivalent of an arithmetic operation to ordering.

Proposals about innate, universal skeletal domains of knowledge need not, and cannot, rest on our ability to show that infants are extremely competent in these domains. The human ability to walk upright is largely dependent on innate contributions, but nobody expects 3- or 4-month-old infants to walk. Alternative accounts of the "baby arithmetic" results (e.g., Simon & Klahr, 1995) are better countered by a cumulative body of converging evidence generated with many different methods. We do not rest our case on the infant results described here.

The Effects of Novelty on Numerical Reasoning

Just such a converging body of evidence is provided by comparisons between how normal and retarded children generate novel solutions to counting problems, together with evidence of cross-cultural structural commonalities in numerical reasoning. The successful ability to invent novel counting strategies implicates at least implicit conceptual understanding of the counting principles, just as the ability to generate novel sentences or make up words implicates implicit understanding of relevant language rules. On the assumption that existing structures encourage children to find successful solutions, indirect hints should suffice to encourage movement toward the generation of a correct solution. But if children lack the requisite conceptual structure, even explicit hints should do little to improve performance (Campioni, Brown, Ferrara, & Bryant, 1984).

Gelman and Cohen's (1988) study of normal and Down Syndrome children's solutions to a novel counting problem illustrates this point. It is well known that retarded persons can have problems with numerical tasks, this being especially so for Down Syndrome individuals (Edgerton, 1967; Thurlow & Turner, 1977). The retarded children's school where Gelman and Cohen's study was conducted took this into account, and incorporated counting opportunities into almost all activities of the school day. Frequency of exposure therefore favored this sample of children over the normal preschool sample in the study. Indeed, nothing in our own observations of preschoolers in a variety of settings suggests that the exposure of normal preschoolers to explicit instruction in counting is as intense and pervasive as it is in the curriculum at the retarded children's school. Still, significant differences favoring the preschoolers were found in the two groups' abilities to count (Gelman, 1982). These differences are consistent with the view that some retarded children do not count with understanding because they lack the core skeletal structure that normal children have. Eight of the Down Syndrome students in the Gelman and Cohen study were poor counters; two were excellent

counters. These differences were perfectly related to the students' ability to invent successful counting solution for Gelman and Gallistel's (1978) novel "doesn't matter" counting task.

Children start the "doesn't matter" counting task by counting a row of heterogeneous items. This done, an experimenter uses a puppet to ask the child to do a counting trick, which is to count all the items but in a way that involves tagging the second item as "the one," or "the three," or "the four," and so on. *Skip-around* and *correspondence-create* are two kinds of solutions children come up with to solve the problem. For example, given a 5-item display, one child pointed to the second item in the row while saying "one," then backed up to point to the first item and say "two," and then skipped over the second item to say "three" while pointing to the third item, and so forth. Another child switched the positions of the first and second item and then simply counted from one end of the row to the other as usual.

The preschool children were more likely to find an acceptable solution as soon as they were given a target problem than were the group of school-aged retarded children who were of comparable mental age but had been scored as poor counters. The preschoolers also benefited extensively from indirect hints, whereas the retarded children failed to benefit at all from either indirect or direct hints. Even when the retarded children were shown an example of a successful trial, they did not transfer on a subsequent trial. Finally, the two groups of children produced different error patterns. The preschool children sometimes rearranged words in the count list instead of objects in the display, for example, "one, three, two, four, five." Although they violated the stable-ordering principle, they at least met the task constraint and ended their count with the correct cardinal value of the display. Down Syndrome children altered their count lists in a different way. When asked, for example, to make the second item "three," they started with that item and said "three" and then continued on in the count list until they tagged all items ("three, four, five, six, seven, eight"). This violated the cardinal principle, an error that is taken by many as evidence for a lack of understanding of the counting principles (e.g., Fuson, 1988).

In short, the various analyses of the different solution efforts of the retarded children who could not count converge on one conclusion. Even though these children received a great deal of counting input, they did not assimilate it with understanding. All indications are that they lacked the implicit counting principles and related arithmetic structures under discussion. In contrast, the data from the preschool

children buttress the hypothesis that they have the implicit principles that outline the relevant domain, even though the relevant inputs they receive are less frequent and picked up on the fly (Gelman, Massey, & McManus, 1991).

Young children's ability to deal with novel number conditions come to the fore again when they are given addition and subtraction problems. In these tasks, they sometimes invent solutions which use counting (e.g., Groen & Resnick, 1977; Siegler & Robinson, 1982; Sophian, 1994; Starkey & Gelman, 1982). This is illustrated in the following protocol from an experiment where children encountered the unexpected effects of the experimenter surreptitiously removing two items from a display of five items (Gelman & Gallistel, 1978, p. 172). It also is an example of the kind of evidence that led Gelman and Gallistel to conclude that young children know that addition increases numerosity and subtraction decreases it, but that they do not know the inverse rule, that is, that the addition of X can be canceled by the subtraction of exactly X.

Child: Must have disappeared.

Experimenter: What?

Child: The other mouses?

...

Experimenter: How many now?

Child: One, two, three.

Experimenter: How many at the beginning of the game?

Child: There was one there, one there, one there, one there, one there.

Experimenter: How many?

Child: Five—this one is three now but before it was five.

Experimenter: What would you need to fix the game?

Child: I'm not really sure because my brother is real big and he could tell.

Experimenter: What do you think he would need?

Child: Well I don't know . . . Some things have to come back.

Experimenter: [Hands the child some objects including four mice.]

Child: [Puts all four mice on the plate.] There. Now there's one, two, three, four, five, six, seven! No . . . I'll take these [points to two] off and we'll see how many.

Child: [Removes one and counts.] One, two, three, four, five, no—one, two, three, four. Uh . . . there were five, right?

Experimenter: Right.

Child: I'll take out this one here [on the table] and then we'll see how many there is now.

Child: [Takes one off and counts.] One, two, three, four, five. Five! Five.

Structural Universality and Cross-Cultural Surface Variability

The structure of such invented solutions is common across tasks (Starkey & Gelman, 1982) and in people from all over the world, be they in schooled environments or not. Although different cultures use different lists and although older individuals might work with numbers in their head and use larger values, the underlying structure of the reasoning is the same. Different count lists all honor the same counting principles (Gelman & Gallistel, 1978) and different numbers are made by adding, subtracting, composing, and decomposing natural numbers that are thought of in terms of counting sets. The mathematical operations involved are always addition and subtraction, even if the task is stated as a multiplication or division one. In the latter case, people use repeated addition and subtraction to achieve an answer. Those who have learned to count by multiples of one, for example, by fives, tens, fifties, hundreds, are at an advantage since they can count and add faster than if they had to count by one (Nunes, Schliemann, & Carraher, 1993; Resnick, 1989; Vergnaud, 1983). This commonality of structure across tasks and settings is an important additional line of evidence for the idea that counting principles and some simple arithmetic principles are universal. So, too, are the cross-cultural studies of "street" arithmetic.

Further evidence for the worldwide use of counting and the arithmetic principles of addition and subtraction comes from studies of "Street Arithmetic." Reed and Lave (1977) found that Liberian tailors who had not been to school solved arithmetic problems by laying out a set of familiar objects (e.g., buttons, pebbles), or drawing lines on paper, and then counting them. Nunes, Schliemann, and Carraher (1993) found that 9- to 15-year-old street vendors were able to indicate what a number of coconuts would cost by performing a chain of additions on known numbers. For example, a 9-year-old said "Forty, eighty, one twenty" when asked the cost for 3 coconuts at a price of 40 cruzeiros

each. When another child was asked to determine how much a customer would have to pay for 15 of an item costing 50 cruzeiros each, he answered, "Fifty, one hundred, one fifty, two hundred, two fifty. (Pause). Two fifty, five hundred, five fifty, six hundred, six fifty, seven hundred, seven fifty" (p. 43).

Variations in performance levels and time to learn the count list are systematically related to schooling, the transparency of a given language's system for generating count words above ten, the degree to which numbers are used in the everyday activities of a culture, and what functions the count list serves (e.g., Gelman & Gallistel, 1978; Miller, 1996; Zaslavsky, 1973; Saxe, 1979). In the past, some conclusions about non-Western inability to count have been due to investigators not knowing to avoid cultural taboos against counting certain classes of things. Others probably followed from their failure to recognize the arithmetization of count-2 systems and/or a failure to recognize that some cultures use hand configurations and body positions as the tagging entries in their count list (Zaslavsky, 1973). It is true that some languages only have two or three count words. It does not follow, however, that the people who use these languages cannot count in a principled way. The Bushmen of South Africa are a case in point. They indeed have but two separate count words. However, this does not keep them from counting at least to 10. They manage the latter by using the operation of addition to generate terms that represent successive larger cardinal values. For example the word for eight translates as "two+two+two+two." This is the very same addition solution that plays so significant a role in "street mathematics." Thus, despite notable differences in the particular verbal counting solutions that different cultures endorse, at the structural level there appears to be a critical overlap.

The Oksapmin of New Guinea use a count list that has 29 unique entries. It starts with the right thumb, which corresponds to "1," the right index finger to "2," and so on (Saxe, 1979). Tagging continues to the right, up through points on the right arm and shoulder, around the outside of the head, down the left shoulder and arm, and ending with the left thumb which corresponds to "29." Should these count lists be rejected as not being "real" count lists because they are "concrete?" No; Saxe (1981) reported that the Oksapmin use their lists in a principled way. Participants in his study were told a story in which people in a far away village count starting on the left side of their bodies instead of the right side. They were then told that men from both villages counted sweet potatoes, ending their counts at

the same body part (e.g., left shoulder). Finally, they were asked who had counted more potatoes. If participants had said that both men counted the same number, we could conclude that they were not using body parts as arbitrary symbols in an ordered list. However, since the participants did answer appropriately, depending on which villager's system was used, we can conclude that they were able to use both counting and reasoning principles.

When the focus is on the abstract level of structure, multiple lines of evidence fit together to support the conclusion that skeletal principles of counting and arithmetic reasoning form a core domain. Similar conclusions are reached by others in their studies of biological classifications (Atran, 1994; Simons & Keil, 1995; Wellman & S. Gelman, Ch. 16, this Volume) and the animate-inanimate distinction (Gelman et al., 1995; Premack & Premack, 1995). This makes it possible for cognitive universals to live alongside culturally-specific interpretations (Boyer, 1995; Gelman & Brenneman, 1994; Sperber, 1994).

The seven postulates described in this section form an interconnected theoretical and empirical framework which any theory of learning and constraints must accommodate. It rests firmly on converging lines of evidence, and does not rely on any one particular study or interpretation. In the next section, we will review different theories of learning and attempt to measure them against this framework.

DIFFERING THEORETICAL PERSPECTIVES ON CONSTRAINTS AND LEARNING

Different theoretical approaches to cognitive development embody different views about the way constraints function in the service of knowledge acquisition and problem solving. Those based on empiricist assumptions about concept acquisition are grounded on the idea that to start, we are limited to processing bits of sensations. This constraint on the nature of first inputs means that inductions about abstract or relational concepts can follow only as a function of the build-up of associative strength. For domain-specific theorists, on the other hand, the principles that organize a domain draw attention to abstract or relationally structured inputs from the outset. Both domain-specific and stage theories share this learning-enabling point of view, whether or not they posit innate skeletal mental structure. These two kinds of theories differ however in that the latter typically focus on broad-based structures of mind. Information processing theorists draw

attention to the existence of limits on our real-time cognitive processes (Shatz, 1978, 1983; Simon, 1972). This, too, is a view of constraints that places limits on the learner. But in contrast with the associationist view, here the limits are about considerations of what learners can do in a given amount of time, not about the learning mechanism itself.

It is not always easy to uniquely identify particular authors as Associationist, Information Processing, Stage, or Domain-Specific theorists. Some assumptions about the role of constraints can be shared across different theories of cognitive development, and some theories contain multiple assumptions about constraints (for a related discussion, see Keil, 1990). The same points apply to domain-general stage theories, as well as approaches including Situated Learning, Cultural Psychology, and Evolutionary Psychology. Many theorists combine two or more approaches. For example, researchers including Case (1992), Halford (1993), and Kuhn, Garcia-Mila, Zohar, and Anderson (1995) combine key Piagetian and information processing notions. Resnick (1994) works with a combination of information processing, situated learning, and domain-specific notions. Parisi (1996) has explicitly set out to develop connectionist models that are consistent with evolutionary and Piagetian approaches. Domain-specific accounts of early cognitive development often do consider maturational constraints (e.g., Johnson & Morton, 1991). Nevertheless, their emphasis is on considerations pertaining to a domain as opposed to general structures of mind or limits on a general processor.

The different usages of the notion of constraint are often joined with qualitatively different usages of other shared terms, especially *learning*. This is because qualitatively different ideas about learning mechanisms are embedded in the foundational principles of the theory. We expand on this point before continuing with a discussion of how different classes of theories deploy the concept of constraint.

About Terms and Their Role in Scientific Theories

Scientific and technical terms are defined separately from nontechnical words in dictionaries, and for good reason: they derive their meaning from the theories within which they are embedded. Technical definitions for terms are often inconsistent with their everyday usage. For example, *multiply* always means *to increase* in everyday use. This is not true in mathematics wherein two fractions multiplied by each other render a smaller number. In the same way, there can be different intended meanings when different scientific theories of learning use the same word. We already have seen that the notion of *constraint* changes in significant ways

as a function of the theoretical context in which it is embedded. So, too, do its semantic brethren *boundary*, *innate*, *biological*, and *instinct*. The everyday meanings of such terms are commonly paraphrased as *restricted*, *required*, *not voluntary*, *not spontaneous*, *hemmed in*, and *forced*. They are pitted against words like *acquired*, *learned*, *educated*, *experienced*, and so on.

This is why we expect some readers to be puzzled by our use of the term "learning" in the context of discussions of innate contributions, since for many, "that is not what learning is really about." The long-standing tendency to paraphrase *innate* as *not-learned* persists in much of psychology. It fits well with the nontechnical definition of *innate* in the *American Heritage Dictionary*: "Of or produced by the mind rather than learned through experience." The same dictionary includes amongst its definitions of *constraint* and *constrain*, "The state of being restricted or confined within prescribed bounds," and "To compel by physical, moral, or circumstantial force; oblige." The broad intuitive appeal of these definitions is as it should be, capturing the terms' everyday meaning-in-use.

It is noteworthy that Hinde and Stevenson-Hinde (1973, p. 470) considered as "unfortunate" the choice of the term "constraint" for the title of the conference that led to their book *Constraints on Learning*. They worried that researchers rooted in the traditions of associationism and learning theory would move to talk about "too-powerful general laws [of association] hedged by constraints," as opposed to "some quite new formulation [that] would seem more profitable." Unfortunately, in many cases, their worry proved prophetic. The alternative they had in mind was a theory of learning founded on the assumption of different domain-specific learning mechanisms, one where the very notion of learning was qualitatively different than one based on associationist assumptions.

It is well to recall that there are many examples in the history of science (e.g., Kitcher, 1982; Kuhn, 1970; Lakatos, 1970) where the meanings of terms differ as a function of the theory about the phenomena to which they refer. Indeed, differences in the intended meaning of terms can be a clue that there either is an ongoing theory change among scientists about the domain in question, or a situation in which qualitatively different scientific theories coexist. Such co-existence characterizes the current state of affairs with respect to the nature of concept learning and development in animals and man (Gelman, 1993; Gallistel, 1995). The same can be said of fields including linguistics (e.g., Chomsky, 1972; Pinker, 1991), comparative cognition and ethology (e.g., Gould & Marler, 1987), and cognitive science (Gold,

1967; Osherson, Stob, & Weinstein, 1986). Given that current assumptions about the "laws of learning" and related mechanisms can vary in fundamental ways, we should be on the lookout for differences in the intended meanings of phrases like *learning theory*, as well as differences in assumptions about the different theoretical premises they entail. A key case in point is how different kinds of learning theorists view the role of innate knowledge.

Anti-nativists who reject the idea that there is anything like innate knowledge typically use what is now an outmoded biological version of the notion, namely some form of genetic determinism. This usage is exacerbated by discussion in the popular and sometimes even scientific press of genes "for" particular traits. In this view, innate contributions are always in a mature steady-state, waiting to generate perfect and therefore nonvariable performance at all times. This perspective is readily recognizable in writings that reject the idea that there are constraints on cognitive and language development. To illustrate, Nelson (1988) wrote, "A true constraint would be manifested in all or none type responses; . . . If the constraint is universal (cognitive or linguistic), all children should follow the pattern. . . . If they are innate, they should apply from the beginning of the language learning process" (pp. 227–228).

Ideas like these about the way our biology makes complex contributions to our cognitive development do not take into account the explosion of work and theory in developmental biology, ethology, animal cognition, behavioral genetics, and evolutionary psychology. Such efforts have led a number of researchers across these and other fields to use the terms *innate* and *learned* in ways that are complementary. Gould and Marler (1987; Marler, 1991) write about the "instinct to learn." Fodor (1983) writes about domain-specific learning modules. Gallistel et al. (1991, p. 30) writes about domain-specific learning engines whose computational processes are suited to the many specific problems the animal confronts:

Within each functionally defined domain of animal endeavor, there can be dramatic differences in the need for flexibility, and thus in the need for learning. There must always be a strong learning component in any mobile organism's ability to develop a representation of the spatial location of objects in the world, as it is extremely implausible that such information is prewired. But in other domains, such as the identification of food or the recognition of conspecifics, species differ as to how much demand on learning their solution to the problem requires. These differences are reflected in the existence and the complexity of specific learning mechanisms.

In each of these latter lines of research, the idea is to work on a kind of learning theory that differs from those in the empiricist tradition. Within this framework, the notion of constraint shifts from one about limits on domain-general learning processes, to one about the enabling effects of domain-specific learning systems and their particular structures. This difference in theoretical perspective has important consequences for how one interprets common terms in the two classes of theories. Some terms can have notably different meanings, just as do the common terms used by children, college students, and physics experts to talk and reason about the nature of heat, light, and electrical current (Gentner & Stevens, 1983; Slotka, Chi, & Joram, 1995; Wiser, 1987). Put differently, the current theoretical environment for discussions about cognitive development and the mechanisms of change is itself dealing with something akin to a theory change. Readers would do well to keep this in mind, especially when the discourse is about learning and constraints.

Constraints in Associationist Theories

Associationist theories of animal learning and human concept formation remain popular (e.g., for treatments of modern versions, see Gluck & Bower, 1988; Kehoe, 1988; McClelland & Rumelhart, 1986; Pearce, 1994; Schwartz & Reisberg, 1991). Modern associationist theories carry forward the key assumptions of Empiricism, including the one that acquisition of knowledge about the world and how to respond to it is derived from the capacity to form associations about sensory data in a lawful way. Initially, the British Empiricists formulated two laws of association: (a) the law of frequency (the more exposures there are to a particular association, the stronger the association); and (b) the law of contiguity (the closer together that the occurrence of an association's components are in time and space, the more likely it is that an association will be formed). Modern versions of the theory add several laws to these original ones. These include: (c) the law of similarity (the more similar the stimuli, responses, and/or existing associations are to each other, the greater the chance that an association will be formed); (d) the law of effect (reinforced pairings and behaviors are more likely to be learned than nonreinforced ones); and (e) the law of contingency (to be learned, the components of an association must be contingently related and not simply contiguous in time and/or space: the first component must predict the second in some sense).

The original theory of association emphasized the principle of equipotentiality: all effective sensory inputs and

all observable responses were treated as equally potential contributors to the associative process. In other words, learning was thought to be a singular, content-independent process. Learning about language, causality, space, time, biology, physics, fishing, number, and other people's minds should therefore be traceable to the same fundamental laws of association that explain a rat's learning to avoid poisonous food, or a pigeon's ability to learn the temporal parameters of a reinforcement schedule. Nothing in classical association theory limits the nature of stimuli or responses that can be associated with each other. Similarly, it provides no principled reason to expect learning to vary as a function of development.

It is probably considerations like these that lead some to think that traditional "learning theory" makes no assumptions about constraints on the learner; that the mind is a completely flexible learning machine. However, association theory does embed one key constraint on the infant learner: young infants are constrained to process and form associations between bits of sensations, or between sensations and reflexes. If their sensory systems are not yet functioning, or their capacity to form associations is weak, then little or no learning can occur. It follows that blind individuals should not be able to learn about the visual world.

These constraints on inputs and associative processing are consistent with classical empiricist dogma: the assumption of an initially blank mind (*tabula rasa*), and the related idea that "there is nothing in the mind that was not first in the senses" (Locke, 1690). However, their implications are strikingly inconsistent with the postulates and supporting evidence reviewed above. Some alternative is needed to encompass the reviewed facts, including ones which show that the mind is actively engaged in its own knowledge acquisition; that some data for learning are supplied by the learners themselves and not from sensory input; that even infants can prefer structured inputs to unstructured bits of sensory experiences; and that cognitive development is nonrandomly selective from the start.

There are different kinds of theoretical responses to these findings of conceptually-driven selectivity, domain-specificity, and non-equipotentiality in early learning. One involves efforts to accommodate such results within the general process theory of association (Schwartz & Reisberg, 1991, Ch. 1). Within modern associationist accounts, constraints act to limit a too-general or too-powerful learning process. A parallel notion of constraint is used in Connectionist theories that assume an associative theory of mind (for a discussion, see Davis & Pérusse, 1988). Most

connectionist accounts limit the construction of knowledge to the build-up of nonsymbolic association networks of knowledge (Bates & Elman, 1993; Elman, 1990; McClelland & Rumelhart, 1986). Here, constraints are added to the model to take into account real-time limits, the nature of human perceptual systems, and so on. However, there are also symbolic connectionist models of learning and cognitive development that build in structural constraints (Parisi, 1996).

The class of accounts of cognitive development which we favor treat the mind as a structure-using and structure-creating device (see below and Karmiloff-Smith, 1992). But even here there are fundamental differences between theorists, especially with regard to the nature and origin of mental structures, and how the notion of learning is construed. Some accounts remain within the associationist tradition, attributing the acquisition of mental structures to the build-up of associative networks (for examples, see T. Simon & Klahr, 1995). Other accounts forge a new path outside the associationist tradition, attributing the acquisition of mental structures to domain-specific information processes that possess specialized functions.

Constraints and Mental Structure I: Information Processing Theories

Among theorists who emphasize the role of mental structures, there are various proposals about how mental structures function as constraints on learning. One class of proposals is information processing theories of cognitive development. Although some information processing theorists include domain-specific knowledge structures in their hypotheses about the architecture of cognition (e.g., McShane, 1991; Mandler, Ch. 6, this Volume), most do not. Still, as McShane points out, there is a shared frame of reference for information processing theorists, this being that the mental architecture is divided into processing components like attention, short-term memory, working memory, and long-term memory. Research topics, not to mention textbook chapter headings, are organized around the idea that there are real-time processing constraints on these cognitive components (Anderson, 1995; Cosmides & Tooby, 1994b).

Information processing accounts of development focus on some limiting effects of an immature brain on the capacities to sense, act, remember, and plan (e.g., Anderson, 1995; Kail & Bisanz, 1992; Newport, 1990; Siegler, 1983, 1991; Simon, 1972). There are straightforward reasons for

the existence of constraints on what infants can process: their sensory system, motoric abilities, and brain functions are all immature. Compared to other primate species, human infants appear to be born early, with brains that continue to develop for months at a rate comparable to that before birth. The cortex in particular develops substantially during the first year of life, particularly in differentiation, myelination, and connectivity (Rabinowicz, 1979).

The upshot of this ongoing development is that immature brain mechanisms may mistakenly give us the impression that infants lack conceptual knowledge. For example, infants 4 months of age and younger are guaranteed to fail any task that requires them to accurately perceive the details of an object across the room. Since their visual system is too immature to process such information (Held, 1993; Johnson & Gilmore, 1996; Kellman, 1996), it is hardly fair to conclude that they lack the conceptual ability required for the task. Likewise, when the stimulus conditions of a task unwittingly elicit a reflexive response that interferes with an infant's overall plan of action, it is inappropriate to reach conclusions about competence (Diamond, 1991). If there are developmental functions on processing capacities, then there will be changes in children's ability to deal with the real-time demands of understanding and performance.

Information processing theorists often pair assumptions of constraints on real-time processing and the knowledge base with hypotheses about how the mind attempts to circumvent these limits. In this vein, development is characterized as involving the increasing release from processing limits, either as a function of maturation, an increasing knowledge base, and/or improved abilities to use organizational strategies (Siegler, 1991; Simon 1981). Information processing approaches to the role of constraints have generated a wealth of interesting research programs, including ones about developments in attentional capacity or strategies (Simon, 1972; Klemmer, 1983), processing space (Case, 1992; Halford & Wilson, 1980), expertise (Chi, 1992), processing speed (Kail & Bisanz, 1992), and strategies for circumventing information processing limits (Bjorklund, this volume; Brown, Bransford, Ferrara, & Campione, 1983; Ornstein & Naus, 1978).

Newport (1990) and her colleagues take a very different tack when they argue that maturational constraints facilitate early learning about the structure of language. They have been working on the intriguing idea that limits young children's information processing capacities actually assist their language learning rather than hindering it. These limits lead children to miss much of the linguistic data in their

environment, and instead to attend to bits and pieces of it. Since such language-relevant data can be characterized as features (for example, the markers "-ed" and "-s" indicating past tense or plurals), the resulting focus on bits and pieces is well suited to the task of inducing the morphemic and syntactic structure of the input language. From this perspective, one might even expect young learners to do better at learning the structure of a language than older ones.

Newport and her colleagues have provided support for this hypothesis (e.g., Johnson & Newport, 1991; Newport, 1990; Newport & Supalla, 1990). They were able to conduct a natural experiment about the effect of starting to learn American Sign Language (ASL) as either a Native learner (born to deaf parents and exposed to ASL from birth), an Early learner (a 4- to 6-year-old whose first encounter with ASL occurred when they started a school for the deaf), or a Late learner (not exposed to ASL until at least the age of 12 and eventually attending a school for the deaf, meeting ASL-speaking deaf friends or a spouse). All participants in the Newport project had at least 30 years of daily exposure to ASL, and all did well on word-order and syntactic phrase parsing tests. Nevertheless, differences favoring the Native learners emerged, having to do with knowledge and ability of how to use the component structure of the morphology and syntax. For example, Late learners had a far larger number of "frozen" signs in their vocabulary, signs that they could not decompose into morphological units. Early learners shared these vocabulary items but could also use them in nonfrozen ways, making up novel words with their components.

Siegler and his collaborators (e.g., Siegler & Crowley, 1994), working within the information-processing frame of reference, have developed an important version of the constraint-as-enabler view, encompassing the conditions that a general purpose central action-assembler has to take into account in order to generate a successful plan of action. Their focus is on the development of competent plans of action, for example, for solving novel arithmetic and tic-tac-toe problems. Siegler and Crowley (1994) propose that 5- and 10-year-olds generate a *goal sketch* that functions to help them generate potentially goal-relevant candidate strategies. Children's ability to judge which of several strategies would be better for a given problem is but one line of evidence presented in support of the proposal. As Siegler and Crowley point out, to accomplish these tasks successfully, it is necessary to select goal-relevant strategies and have some way to evaluate whether they work.

Closely related arguments are found in Gallistel (1985) and Gelman and Greeno (1989), even though both are domain-specific theorists. Gallistel's emphasis is on the way domain-specific goals potentiate and depotentiate relevant action components to assemble functionally coherent behavioral sequences. Gelman and Greeno's effort is focused on how numerical *conceptual competence* in conjunction with *interpretative competence* relates to the ability to generate competent plans of action for a given task and setting.

All of these lines of research highlight the fact that the ability to create a competent plan of action requires much more than a general problem-solving facility: it requires a solution to what has been called the "frame problem." For instance, Newell, Simon, and Shaw's (1958) General Problem Solver model does not on its own have the capacity to assemble task-relevant and content-relevant solutions. As such, it could go on forever making up solutions that have nothing to do with a non-assigned problem. Siegler and Crowley's goal sketch helps solve the frame problem by virtue of its ability to search existing knowledge bases that might be relevant to the task at hand, and also to use the search results to put together a plausible strategy (at least some of the time). The result is a planning process that benefits from enabling constraints.

Constraints and Mental Structure II: Stage Theories of Cognitive Development

Within stage theories, structures of mind are typically described with reference to single structures or formats that are domain-general and content-independent. For example, the Piagetian stage of Concrete Operations encompasses the ability to identify, learn, and reason about well-classified data, various kinds of orderings, and number. Domain-specific theories instead postulate multiple structures with domain-specific organizing principles, for instance about number, animate natural kinds, and so forth. The learner's ability or inability to find relevant inputs is due to the availability or non-availability of mental structures that can be mapped to domain-relevant data.

Piaget's Concrete Operations are well-known for their role as mental structures that enable the ability to organize data into hierarchical classification schemes and reach inductions about the inclusion relationships between superordinate and subordinate levels of a given classification (Piaget & Inhelder, 1956). Within the theory, preoperational children have yet to achieve operational structures

and therefore are constrained to make classifications on the basis of surface, perceptual similarities. This leads them to solutions that count as failures on a wide range of classification tasks. For example, they respond to class-inclusion questions on the basis of the perceptual data given. When shown a display of ten roses and two tulips, and asked, "Which is more, the flowers or the roses," they chose the roses, as if they interpreted the question as, "Which has more, the bunch of roses or the bunch of tulips?" (Piaget & Inhelder, 1956).

There are many challenges to Piaget's theory and findings about classification. Some of these are taken up below in our discussion of the nature of concepts. At the moment however, we focus on a key feature of stage theories. Stages are structures of mind that not only limit but also enable cognition and learning. For example, sensorimotor schemes constrain infants to learn about sensory and motoric representations, pre-operational thought constrains children to learn about the perceptual aspects of their world, and so on. But stages also have a facilitating function, especially when they achieve the status of a logical structure. For Piaget, concrete operations underlie several abilities: to think about sets of objects in terms of the dimensions that characterize them, to find the intersecting values of two dimensions, and to reach inferences about members of a class-inclusion hierarchy. Similarly, logical structures enable the use and understanding of classifications, seriations, and numerical concepts (Inhelder & Piaget, 1956; Piaget, 1971). When children have developed Concrete Operations, they can use them to organize a seemingly disparate set of objects into systematic subclasses.

Constraints and Mental Structure III: Domain-Specific Theories of Concept Learning

A shared assumption across stage- and domain-specific theories of mind is that mental structures find, interpret, and assimilate data whose structure they overlap with. For this reason, domain-specific knowledge structures, even when they are skeletal, can enable learning by helping individuals find inputs that have the potential to nurture further learning about a domain. Because domain-specific structures lead learners to select one class of inputs over others, one might say they also have a limiting role. But in this case, the limiting role is due to their ability to select from the vast flow of possible data those inputs that are relevant for learning about a domain. It is not that learners cannot learn about other things, but rather that different

mental structures support different kinds of learning. In all cases, the mental structures actively seek out data that are structurally consistent with the principles that organize them. The result is that learners are more likely to encounter domain-relevant learning paths and the consequent benefits that we discuss below.

Among domain-specific theorists, some assume that there are innate skeletal learning systems for core domains. Domain-specific theorists also differ on whether infants' core domains are already content-rich enough to be called theories. Whereas Spelke (Carey & Spelke, 1994) is willing to grant infants an implicit theory of physical objects, we prefer to limit the attribution to the different skeletal principles organizing the causes of movement and change of animate and inanimate objects (Gelman et al., 1995). Although we prefer the view that core domains start as different sets of skeletal principles, there are others who refer to such domains as being organized by core theories (e.g., Carey & Spelke, 1994; Gopnik & Wellman, 1994; Wellman & S. Gelman, Ch. 11, this Volume). To the extent that theories are organized by a set of domain-specific principles, they certainly meet our definition of a domain (Gelman, 1993; and below).

The open question is how much infants already know about the content of a domain and therefore whether it is best to assume their knowledge is skeletal in form. In either case, the move to postulate core domains paired with an assumption of domain-specific learning systems was much influenced by students of animal learning who work within an evolutionary perspective. These lines of influence are illustrated below in our development of the idea that constraints often serve as enablers of learning, that the mind will use existing knowledge structures to actively learn more about a particular conceptual domain.

Our preference for a domain-specific structural theory of concept learning is reinforced by the need to deal with the well-known empirical and theoretical difficulties with stage theories of classification abilities (Gelman & Bailargeon, 1983). Since stage theories of classification are grounded in the classical theory of concepts, these difficulties are shared with the classical (Empiricist) account of concepts. This account assumes concepts are built up from a primitive set of features available to the mind. Like other cognitive developmentalists and cognitive scientists, we believe that the difficulties faced by the classical theory warrant adopting the position that the meaning of a concept is embedded within the structure of the domain to which it belongs. This is one of the main reasons for our view that

enabling constraints privilege some learning over others and are structural and conceptual in kind. Examples of conceptual constraints include Markman's (1989) mutual-exclusivity constraint on the labels for items in a common classification structure (for considerations pro and con of word-learning constraints, see Markman, Ch. 8, this Volume; Keil's [1989] M-constraint on the members of ontological categories; or Gelman & Gallistel's [1978] constraint on double-counting embedded in the one-to-one counting principle).

Summary

To recapitulate our review of differing perspectives on constraints, the term *constraint* does not stand alone. As shown in Table 12.1, its meaning varies as a function of the theory in which it functions, as well as each theory's foundational assumptions about mental architecture and knowledge acquisition mechanisms. Given a domain-specific epigenetic theory of concept acquisition, enabling constraints serve the learner in significant ways. However, this claim is not tied to the assumption that a particular domain of knowledge is a core domain which develops from innate skeletal principles. In the remainder of the chapter, we further develop the idea of enabling constraints. It is not the presence of domain-relevant conceptual constraints that stands in the way of learning and problem solving; rather it is their absence. Readers should notice the overlap between this conclusion and the one above about Siegler and Crowley's (1994) concept of a goal sketch. The account of cognitive development that we will articulate in the remainder of the chapter introduces the idea that mental learning tools are recruited along with domain-specific knowledge as facilitators of learning and problem solving.

ENABLING CONSTRAINTS, EPIGENESIS, AND DOMAIN SPECIFICITY IN LEARNING

A successful theory of concept acquisition and cognitive development must address and accommodate all of the following issues: (a) the sources of both initial and subsequent knowledge structures; (b) the ways those structures foster learning; (c) the characterization of relevant inputs for both early and subsequent learning; (d) the characterization of the sources of those structured inputs, including the physical and sociocultural environment; (e) the nature of the individual's active participation in the elaboration of

TABLE 12.1 Summary of the Relationship between Theories of Concept Acquisition and Their Notions of Constraint

Framework	Kinds of Constraints	Examples/Implications
I. Traditional Associationism	<ol style="list-style-type: none"> 1. Learning is governed by the laws of association: frequency and spatial/temporal contiguity. 2. Nothing is in the mind that is not first in the senses. (John Locke) 3. Concepts are built from sensory primitives. 	<ul style="list-style-type: none"> • Only stimuli that fall within the associative window (closely paired in time and space) can be associated. • Blind children should not be able to learn the meaning of "see," because they have no experience of seeing. • Young infants do not have concepts.
II. Biological Boundaries on Associationism	<ol style="list-style-type: none"> 1. Same implications as above, except the parameters of the associative mechanism may differ from domain to domain. Some knowledge is <i>implicit</i> in these parameter settings. 	<ul style="list-style-type: none"> • Taste is more readily associated with illness than is visual appearance. • Taste-illness associations can develop even when illness occurs hours after the taste. • <i>Implicit</i> knowledge in the above two parameter settings: the fact that poisons act by virtue of their chemical composition (which is better indicated by taste than vision), and that poisons may act slowly.
III. Connectionism* (Nonsymbolic)	<ol style="list-style-type: none"> 1. Much the same as traditional Associationism, plus the Delta Rule (the Rescorla-Wagner law). 	<ul style="list-style-type: none"> • Acquisition of the past tense "rule" is due to amount of training, not a language-specific domain.
IV. Information Processing	<ol style="list-style-type: none"> 1. Domain-general information processing mechanisms make it possible to extract structure from input. 2. Real-time constraints on domain-general information processing mechanisms limit what can be learned. 3. In some cases, these limits may simplify the learning task. 4. However, nothing in these general-purpose mechanisms endows the learned representation with any explicit structure that was not in the sensory input to begin with. 	<ul style="list-style-type: none"> • Chunking makes it possible to organize information into learnable chunks. • Short term memory limits what can be learned without considerable effort, e.g., complex syntactic constructions, or long non-rule-governed count lists. • Newport's (1990) "less is more" account of language acquisition. • Rules and principles must be induced, e.g., the counting principles.
V. Stage Theories	<ol style="list-style-type: none"> 1. The structure of one's stage limits what can be attended to and learned in that stage. 2. The assimilation of experience can, however, lead to a metamorphosis, a change to a new structure that enables successful solutions to a broad set of problems. 	<ul style="list-style-type: none"> • A child can assimilate only those inputs that fit their level of mental structure. • The cumulative impact of inputs that don't fit causes accommodation of the child's mental structure. Format change during Concrete Operations supports ability to form abstract concepts.
VI. Domain Specific Learning Mechanisms	<ol style="list-style-type: none"> 1. Skeletal learning mechanisms enable active attention to and processing of domain-relevant data. Different core domains benefit from different learning mechanisms. 2. Some of what is <i>explicitly</i> represented may derive from the domain-specific information processing mechanism. 3. Existing structures may interfere with learning of expert knowledge if they do not map to the structure of the expert knowledge. 	<ul style="list-style-type: none"> • Mechanisms for song learning make it possible for birds to learn their species' song but do not contribute to the learning of counting, a map of the environment, or other domains. • Language learning mechanisms make language learning possible. • First principles of counting and arithmetic support acquisition of natural number concepts. • Space (spatial relations) is explicitly represented as three-dimensional even though all visual input is two-dimensional and all auditory input lacks a spatial dimension entirely. • Newtonian mechanics, calculus, chess, etc., are noncore domains and require learning new principles. Mastery is hard, requiring time and considerable practice.

TABLE 12.1 (Continued)

Framework	Kinds of Constraints	Examples/Implications
VII. Ecological Constraints	1. The acquisition mechanism itself is not specified, because the environment has such a rich structure that it dictates the structure of what will be learned.	• Affordances of the kind addressed in the Gibsons' theory (e.g., J. Gibson, 1983), such as learning about the difference between icy and smooth surfaces.
VIII. Situated Cognition	1. Version of VII in which it is argued that what is learned is specific to the situation or context in which it is learned; hence, there is no cross-situational transfer.	• Street Math does not transfer to School Math.

*The last few years have witnessed rapid changes in the nature of Connectionist models, including examples of domain-specific structural models (e.g., Marcus, 1995), or even a combination of associationism and structuralism.

their knowledge structures; (f) the mechanisms of change which relate initial learning to later learning; (g) the variability in performance levels during learning, across tasks and cultures; and (h) the existence of cross-cultural universals of cognition. A tall order, to be sure.

Most children in nondeprived learning environments acquire core knowledge bases and a conceptual worldview which is shared with the rest of their community if not beyond. They do this despite facing a great deal of ambiguity during their formative early learning years. We do not deny that this commonality of outcome surely must be nurtured by relevant commonalities in learning environments. Again, however, environments alone are insufficient. How do novice learners of any age find these relevant commonalities, given that environments are variously ambiguous and undetermined? How do learners know what is relevant, given that there is an unending flow of possible inputs? How does the mind select what subset of the world to attend to, interpret, and learn about?

In the foregoing sections, we have begun to lay out our reasons for favoring a rationalist-constructivist theory of learning and cognitive development. We believe such a theory to have the best hope for answering these questions, and also for meeting the desiderata listed above. Our view emphasizes the role that domain-specific constraints play in enabling early learning and cognitive development within core domains. The mind starts out with certain core skeletal domains of knowledge that include the ability to actively engage with and benefit from potential relevant environments. Enabling constraints guide learning by allowing the mind to actively and selectively seek out domain-relevant information, and by privileging particular interpretations of that often ambiguous information. They capture universal structural principles about core domains, yet they also lead to variability in learning outcomes due to

epigenetic interaction with particular local environments. Many of these structural principles are abstract, encoding notions such as number, causality, force, and mental states which are not directly perceivable by the senses.

Knowledge growth in core domains comes about as a function of an epigenetic, reciprocal interaction between multiple sources of structured information. Key players in the epigenesis of knowledge include the following: (a) existing domain-specific knowledge structures, be they skeletal or not; (b) the surrounding and internal physical and biological environments; (c) the social and cultural milieu; and (d) the domain-relevant data that the mind itself generates and sets up as it acquires a knowledge base.

We turn now to expand on our view of enabling constraints and epigenesis, and we begin by discussing an evolutionary perspective on learning. This perspective provides a line of converging theory and evidence that meshes well with our view and with the postulates described at the beginning of the chapter. In addition, it fits poorly with traditional associationist views of learning. We then attempt a more detailed development of what enabling constraints are and how they function, followed by a discussion of difficulties faced by the alternative views of constraints reviewed in the previous section. Because of our commitment to learning as an epigenetic mind-environment interaction, we turn next to a characterization of the structured environments in which learning takes place. Last, we address the important issue of conceptual learning in noncore domains. We introduce the idea of mental learning tools, including structure mapping and frequency computing, as mechanisms which operate both in core and noncore domains. They, too, "enable" learning, but in a different way from enabling constraints: they provide bridges between early and late learning, and between domain specificity and domain generality.

Constraints as Learning Enablers: An Evolutionary Perspective

The brain is no less a product of natural selection than the rest of our body's structures and functions. The mind, like all other biological organs, puts its structures to work to accomplish their functions once prerequisite sensory, motoric, and brain maturation levels are reached. Some of these mental structures serve regulatory purposes, such as temperature maintenance, while others involve learning. The function of this latter type of structure is to find, assimilate, and use relevant structured information in order to learn about particular domains and solve domain-related problems. Complex adaptive functional organization is the hallmark of natural selection (Dawkins, 1995), and is a feature of mental processes such as learning just as it is of nonmental processes such as respiration. Complex structures evolve to fulfill a function, to enable the implementation of an adaptive process. Hearts evolved to pump blood, livers evolved to extract toxins, and some mental structures evolved to enable the learning of certain types of information that are necessary for adaptive behavior.

As was the case with sociobiology in the 1970s, current efforts to frame psychology within evolutionary theory (sometimes referred to as "evolutionary psychology") have been criticized for "just-so" storytelling. Debate over the merits of the "adaptationist program" has pointed out the difficulties of identifying adaptations (e.g., Daly & Wilson, 1995; Gould, 1991; Gould & Lewontin, 1979; Mayr, 1983; Piattelli-Palmarini, 1989; Symons, 1990, 1992). These difficulties particularly beset the search for mental adaptations, whose functional structure cannot be directly perceived under the microscope. In addition, evolutionary psychology as a "grand theory" (e.g., Tooby & Cosmides, 1992) must beware of overreaching its applicability throughout the cognitive sciences, and also of underemphasizing converging theoretical and empirical contributions from sources which do not draw on evolutionary theory. Premack (1996) summarizes many of these cautions in his review of the research programs described in Barkow, Cosmides, and Tooby's (1992) book *The adapted mind*. Notwithstanding these potential pitfalls, the adaptive-evolutionary approach provides important insights into the role of constraints in enabling learning (see also Rozin, 1976; Rozin & Schull, 1988). These insights converge with formal theoretical work on learnability (e.g., Chomsky, 1965), as well as empirical findings from learning and cognition in young children, infants, and animals (see Postulates 4 and 7 above).

From an evolutionary perspective, learning cannot be a process of arbitrary and completely flexible knowledge acquisition. In core domains, learning processes are the means to functionally defined ends: acquiring and storing the particular sorts of relevant information which are necessary for solving particular problems. An important class of these problems involves making predictions. These predictions can be about matters such as the future actions of friends and foes; the prospects of other people as mates, allies, or prey; or the likelihood of finding food at a particular location. Predictive strength is much more important than simple associative strength: compared to a vast number of arbitrary correlations, there are relatively few predictive/causal links in the world. As described above, behavior theory has moved in this direction, emphasizing contingency over simple contiguity, partly in response to empirical evidence of selectivity in associations (e.g., Garcia & Koelling, 1966).

The associationist proposal that all types of information are equally learnable (equipotentiality) is inconsistent with the evolutionary view of learning as a set of complex functional solutions to individual adaptive problems (Rozin & Schull, 1988). The mind does not learn all different sorts of information with equal facility, because different sorts of problem solving require different sorts of information about the world to be learned (Gallistel, 1990). More significantly, we should no longer think of information as entirely "out there in the environment," isolated from the information-processing problem(s) whose solution requires that it be learned. There are many types of adaptive information-processing problems: object perception, object classification, navigation, and mate selection to name but a few. Accordingly, there are many different types of learning. Core domains encompass those adaptive problem-solving behaviors which are universal, in which learning proceeds rapidly and without formal instruction.

Enabling constraints embodied in mental structures that support core domains solve the two deep obstacles to learning discussed above: convergent/divergent ambiguity, and the need for guided (nonrandom) selectivity. They selectively guide the learner's attention to relevant inputs, and they selectively guide the learner's interpretations of those inputs toward accurate and adaptively useful ends. These twin functions enable the learner to travel a successful learning path. We wish to emphasize here our use of "a" successful learning path rather than "the" successful learning path: enabling constraints do not inflexibly require an

exact set of inputs to be matched, and neither do they force all learners to learn exactly the same thing.

The evolutionary history of our species has made us good learners in certain problem-relevant areas and poor learners in other areas. Other species of animals may be good or poor learners in a different set of areas, owing to differences in their problem-solving needs (Gallistel, 1990). Human specialties include visual pattern recognition, navigation, and language, while humans are completely outclassed by the location memory of certain nut-hiding birds (Vander Wall, 1982) and the taste-learning abilities of the white rat (Rozin & Kalat, 1971). The fields of cognitive psychology and artificial intelligence have traditionally focused on the flexibility and generality of human learning abilities, usually in the "higher learning" sense that (for instance) a college student can choose to study any academic field. Yet this apparent flexibility should not obscure underlying species-specific conceptual learning mechanisms which are specifically tuned to certain types of information.

All adaptations, both mental and non-mental, have an informational component: they embody information about the environment in which they evolved. Information is at the heart of the well-accepted biological concept of adaptiveness in an ecological niche:

As Young (1957), Lorenz (1966), and others have emphasized, we recognize adaptedness as an informational match between organism and environment. An animal that is well adapted to its environment can be regarded as embodying information about its environment, in the way that a key embodies information about the lock that it is built to undo. A camouflaged animal has been said to carry a picture of its environment on its back. (Dawkins, 1982, p. 173)

If core domain structures of mind were not related to data structures in the world, infants would be at risk for missing evolutionarily important inputs (the selectivity problem) or making up representations of non-existent worlds (the ambiguity problem). In core domains, structural regularities which carve the physical learning environment into different domains have been internalized into the cognitive architecture of the mind (cf., Shepard, 1994). This allows children to focus on domain-relevant aspects of data, and to recognize the often abstract relationships between entities in a domain.

Having evolved to function successfully in a certain environment, like a key, adaptations may or may not perform

as successfully in a different environment. Sperber's discussion of relevance (Sperber, 1994; Wilson & Sperber, 1986) is especially helpful here. He distinguishes between the "proper" and "actual" domains of a cognitive module. Learning mechanisms evolved to serve specific functions: to attend to and store particular sorts of information. The proper domain of a learning mechanism consists of the information that it is functionally designed to learn, that it is in some sense "expecting." The actual domain is whatever structured information in the modern world and in a particular cultural setting is able to engage the learning mechanism: a new lock for an old key. This distinction between old solutions and new problems undercuts a common misconceived criticism of the application of evolutionary theory to psychology. It is often claimed that evolutionary theory expects all modern behavior to be adaptive or even "optimal" since it has survived the process of natural selection. Of the many reasons why this is not so, the most striking is that modern environments can differ dramatically from the environments in which human beings evolved as hunter-gatherers. Adaptations well-suited to the "environment of evolutionary adaptedness," be they mental or non-mental, sometimes fit less well in modern environments. For example, our lungs perform their function poorly in the oxygen-poor environments found at extreme altitudes or in polluted cities.

Returning to enabling constraints as information-processing adaptations, maladaptive learning and performance can occur when there are mismatches between the expected and actual information coming from the environment. Cognitive mechanisms may produce incorrect solutions when provided with inputs whose structure is unexpected. For example, shape detectors in the human visual system appear to "assume" an overhead direction of illumination, and can be fooled in indoor settings where artificial lighting comes from below (Ramachandran, 1988). In general, learning in a noncore domain can be either facilitated or hindered by the degree to which it is consistent or inconsistent with a core learning mechanism that has evolved to selectively "expect" particular types of structured information (Slotnick & Chi, 1996). Relatedly, mental structures which evolved to serve one function may take on additional functions when presented with modern inputs whose structure overlaps with that of the inputs they are expecting; Piattelli-Palmarini (1989) and Gould (1991) describe these modern functions as "exaptations." We take up these issues of transfer in more detail below.

What Enabling Constraints Are and Are Not

Constraints in Core Domains: Defining Relevance and Structuring Early Representations

Stage theories and the classical theory of concepts both run into the same problem: neither has an a priori rule of salience (see Postulate 3 above). What leads us to think that color is usually a relevant feature of a tomato but not a car? The oft-favored rule of salience, that surface features are most relevant to our concepts about objects, turns out to fail under a variety of conditions. If 3- and 4-year-old children are asked to name red foods to take on a picnic, they will include watermelon alongside tomato, ketchup, and plums, even though the outside of a watermelon is green (Macario, 1991). Although color has some predictive value for our ability to identify foods, color does not have good predictive value for identifying cars (save perhaps for some subordinate-level cars with special functions). Similarly, if adults or 3- and 4-year-old children are asked to select or group pictures of items that can go up and down a hill by themselves, their choices do not appear to follow simple similarity. They pick out an echidna, a tarantula, a child, a sloth, and even a lobster, but they do not select statues or dolls—even though these share human parts such as faces, legs, and arms (Massey, 1989).

The difference in the predictive value of color for identifying instances of the categories of cars versus foods has led some to adopt a theory of concepts that assigns different weights to the same perceptual features for different concepts. We could say that the weight for color approaches 0 given the concept of car, and approaches 1 given the concept of food. But this approach will not do either. As soon as we start to assign different weights to the exact same attributes as a function of concept type, we slip in a circular decision rule that embeds in it an understanding of the very thing that is to be understood. In other words, the weighting method assumes that we use our different concepts to decide what weights to assign to the seemingly “same” features. If so, then they cannot be independent, content-free, decomposable intensional features as required by classical learning theory. We therefore are back to square one with regard to the question of what leads us to decide that a perceptual attribute is more or less important, or more or less similar to ones already learned about.

To restate the problem, we cannot avoid a basic question when asking how children learn a concept (or anyone does for that matter): how should we define relevance or salience? Some argue that the problem is solved for us by others, because we attend to what others bring to our atten-

tion (Tomasello, 1992). There is no question that if we did not use such information, matters would be even worse (Leslie, 1995). Nevertheless, how do different individuals converge on a common interpretation of what to both look at? Quine's (1960) “gavagai” problem still remains: each individual could be attending to any of a very large number of aspects of that part of the world within view. For example, an adult might be looking at the fur of a cat while a child is looking at its moving tail. What makes it more likely that a novice will attend to what another person has in mind? How do blind children develop visual concepts as they do (Landau & Gleitman, 1985), including the idea that sighted people cannot see through barriers?

Given the facts of selective attention and ambiguity (Postulates 1 and 3), we can assume that initial choices of inputs and interpretation are made either randomly or in a principled manner. For us, the existence of shared, early, on-the-fly, universal learnings argue against the random selection alternative. So, too, does the fact that infants and animals actively attend to domain-relevant structured inputs, rather than focusing on bits of light, sound, pressure, smells, and so on.

These observations fit with the evolutionary perspective discussed above, that treats core domain learning mechanisms as functional information-processing adaptations that evolved to use particular classes of data to solve particular problems. The determination of what is perceptually relevant is a function of the learner's current domain-specific goals. If she is concerned about classifying different birds, then she is more likely to attend to color. If she wants to figure out how to move a given object, she had better attend to its size and weight. But if the aim is to count the number of objects in a room, she can ignore attributes such as size, weight, and color. This is why we conclude that domain-specific knowledge structures help us determine what kinds of inputs are relevant to their use. Whether perceptual data are relevant or not depends on the conceptual goal we set.

If we grant infants skeletal domain-specific structures of mind, we can develop an account of which concepts might and might not be universally shared. These structures encourage learners to move onto learning paths that will lead them to the kinds of domain-relevant structured environments that can nurture the build up of shared knowledge structures. Assuming that the learning path is strewn with cases of potential inputs that share the structure of the core domain—be these inputs offered by knowledgeable members of the community, developed by the child, or part of the artifact environment—progress is highly likely given

the learner's inclinations to use existing mental structures when interacting with the environment. We emphasize that there must be such domain-relevant learning paths: otherwise, domain-specific, innate contributors to knowledge acquisition will remain skeletal or even wither. On the one hand, were there no skeletons to dictate the shape and contents of the bodies of pertinent knowledge, then the acquired representations would not cohere. On the other hand, were there no supporting environments for learning to interact with, influence, and be influenced by, then epigenesis could not take place.

More on the Notions of Domain and Principle¹

We define a domain of knowledge in much the same way that formalists do, by appealing to the notion of a set of interrelated principles. A domain's principles and the rules of their application serve to identify the entities belonging to that domain, and exclude those entities which do not. Since different structures are defined by different sets of principles, we can say that a body of knowledge constitutes a domain of knowledge to the extent that we can show that a set of interrelated principles organize the entities and related knowledge as well as the rules of operation on these. Principles "carve the psychological world at its joints," producing distinctions which guide our differential reasoning about entities in one domain versus another. In this way, domain-specific structures encourage attention to inputs that have a privileged status because they have the potential to nurture learning about a domain; they help learners find inputs that are relevant for knowledge acquisition and problem solving within the domain.

General processes like discrimination, or general processing mechanisms like short-term memory, do not constitute domains. Nor does a script structure constitute a domain. Scripts are analogous to the heuristic prescriptions for solving mathematical problems, which should not be confused with the mathematical domains themselves (algebra, calculus, theory of functions, and so on). Still, information processing limits on short-term memory can influence whether a given domain-specific problem is solved correctly or not. For example, although there is nothing in Gelman and Gallistel's (1978) counting principles that requires children to place items in a row and count from one end to the other, efforts to honor the one-to-one principle favor this kind of a solution. In order to count without error, they should not double-count or leave any items out. If they

generate an orderly plan of action, we can say that they have a well-formed plan for solving the task.

There is nothing about our definition of a domain that limits domains to those that have an innate basis, or that requires all principle-based, domain-specific knowledge be built on an innate foundation. For instance, dog show judges' knowledge of various species of dogs constitutes a noncore domain of knowledge, rather than an innate core domain (Carey, 1996). In this regard, our insistence on the distinction between core and non-core domains helps deal with a key worry about domain-specific accounts of the mind. As Leslie (1994b) put it: "The trouble with the notion of domain specificity is that there could turn out to be too many domains" (p. 120). We concur with the suggestion that the number of innate domains is likely to be limited to ones that are universally shared, and not dependent on access to formal schooling (a very recent invention). There are few Leonardo DaVincis who acquire a remarkable number of noncore domains of knowledge, but there are some. More typically, different people acquire expertise in different noncore domains as a function of their interests and their commitment to the extended learning effort required for achieving principled understanding.

Only core domains benefit from what we call skeletal knowledge structures. Leslie's (1994b, 1995) idea of "core cognitive architecture" refers to "those human information-processing systems that form the *basis* for cognitive development *rather than its outcome*" (emphasis added). If children are able to interact within the bounds of the everyday environments of their communities, they will acquire knowledge about core domains early on, and with relative ease. Not only do the principles of a domain help children find relevant inputs, they also collect and keep together assimilated pieces of relevant information. In a sense, they are like memory files for storing pieces of domain-relevant knowledge together. This allows for the collection of a database, one that eventually can become large enough to warrant review, support inductive reasoning, and perhaps be reorganized (cf., Karmiloff-Smith, 1992).

Noncore, nonprivileged domains of knowledge are typically acquired later—from middle childhood on into adulthood—and usually are dependent on structured lesson plans and/or focused study. Although not impossible, it is difficult to acquire these new conceptual structures. As discussed above, their acquisition requires the concurrent acquisition of both new organizing principles and a coherent body of relevant knowledge governed by them. It is not surprising that learning about noncore domains generally is effortful and time-consuming (e.g., Anderson, 1995). Core

¹ This section was based on Gelman 1991, 1993.

domains of knowledge can sometimes be barriers to the later acquisition of noncore domain-specific learning (Hartnett & Gelman, in press; Slotta, Chi, & Joram, 1995; see also our discussion of structure mapping at the end of the chapter).

In sum, when we suspect something is a conceptual domain, we can test whether it is by seeing whether it is possible to characterize it in terms of a coherent set of operating principles and their related entities. This definition of a knowledge domain is neutral about the origin of the domain: it can start as an innate skeletal set of principles or be learned from scratch. There is no necessary link between knowledge's domain specificity and its time of acquisition; we should not conclude that early learning is domain-specific and later learning is not.

Constraints, Epigenesis, Universals, and Variability

The classes of relevant inputs for a domain are defined in terms of the structures of that domain, not particular surface features. The actual examples of inputs to learners can differ from one time to another and from one culture to another. This means that although different cultures can offer their young different subsets of domain-relevant inputs, the children will develop concepts that share common structural cores across cultures. Universal innate concept learning mechanisms do not rule out culture-specific knowledge for a given domain, however. Surface differences, for instance in languages or local wildlife, need not alter structural regularities that can still be learned with different exemplars. Just as children in different language environments share knowledge about linguistic universals, so, too, children who encounter different kinds of toys, artifacts, food, physical environments, and so on, will share knowledge about counting principles, the key differences between animate and inanimate objects, and so on. All will encounter animate and inanimate objects, separably moveable objects, collections of things, and so on. The upshot in many cases is that children in different cultures assimilate different domain-relevant knowledge to their shared skeletal principles.

The above examples demonstrate one source of variability that is consistent with a domain-specific account of core domains of knowledge. Another is variability due to limits or differences in knowledge of conversational rules and task requirements, opportunities to practice, planning abilities, and levels of knowledge. Siegal (1991, 1996) presents a wealth of evidence on the way children's limited knowledge of conversational rules can mask their ability to

display conceptual competence. Some might conclude that these sources of variation implicate a situation-specific model of cognition (e.g., Lave & Wenger, 1991). However, there is an alternative to the situated cognition approach, and this is to take on the task of uncovering and describing the sources of systematic variation and how they interact with a model that assumes conceptual competence (Gelman, 1978).

Gelman and Greeno (1989) offer one example of this approach. They identify the understanding of conversational and task variables with the development of interpretive competence, and detail how the acquisition of levels of knowledge is related to the growth of conceptual competence. Recent studies by Kuhn (e.g., Kuhn et al., 1995) and Siegler (e.g., Siegler & Shipley, 1995) provide additional examples of how to seek careful accounts of the kinds of sources that make systematic contributions to variable performance. Researchers can and should return to the problems presented by competence-performance distinctions across development (see Gelman, 1993, for further discussion).

Problems with Alternative Views of Constraints

Associationist Constraints and Theory: Empirical Difficulties

The Postulates section at the start of this chapter reviews some remarkable results which fit very poorly with traditional theories of learning and constraints. "Screening-out" effects, in which young children and animals prefer geometrically ambiguous information over seemingly salient contingent information, are hardly what one might expect if committed to an associationist theory of mind. Relatedly, the many demonstrations of infants' active use of organized perceptual and conceptual inputs mean that we can no longer assume that infants wait passively for bits of sensory data to impinge on their sensoria. They seek out and structure the material that they learn.

How therefore should we characterize the raw data that a learner can learn about (more on this topic below)? This question helps make explicit a theme that is woven throughout this chapter. When the learner was thought to be passive, the "objective" structure of data "in the environment" might be imagined to determine the structure of what was learned. But when the learner helps structure the data, it is much more difficult to know how best to characterize it from a psychological point of view. We continue our discussion of this theme by focusing on the fact that the great

stage theories of cognitive development carry with them a commitment to the classical theory of concepts, a theory that is grounded in the Associationist account of mind.

Domain-General Stage Theories and the Classical Theory of Concepts: Empirical Difficulties

Whatever the differences between Bruner's, Piaget's, or Vygotsky's theories of cognitive development, they all converge on the conclusion that there is an ordered shift from a stage where children use non-hierarchical rules to one where they have the capacity to represent hierarchical classification systems. Therefore, the form of young children's concepts differs qualitatively from those of older children and their elders. They can think about complexes (Bruner, Greenfield, Olver, et al., 1966), pseudoconcepts (Vygotsky, 1962), or precepts (Inhelder & Piaget, 1956), but not about "true" concepts. Two kinds of evidence challenge this stage view of classification and cognitive development. These include demonstrations of young children's success on classification tasks, and penetrating challenges to the classical theory of concepts with its reliance on the mental ability to operate logically with classes. Fodor (1972) points out that traditional stage theories of concept development are grounded on the classical theory of concepts. His argument continues that since the classical theory is deeply flawed, so, too, are traditional stage characterizations of cognitive development.

Domain-Specific Classification Effects. Young children's failure on a variety of classification tasks is well-documented (Inhelder & Piaget, 1964). They concluded that there are three stages in the development of classification abilities. Initially (around 2 to 5½ years) children produce "graphic" solutions, in which items are grouped for no apparent reason or on the basis of a shifting classification criterion. Examples include a train and a cake, or a sequence of objects which taken pairwise have overlapping properties, such as red square, red circle, yellow circle, yellow triangle, and so forth. In the second stage (5 to 8 years), "nongraphic" solutions dominate that are more mature. At the end of this stage, children can consistently use a single criterion for classifying a set of objects, and can then reclassify the objects using a different criteria. Despite this advance, children who generate such solutions are still considered pre-operational when they fail on class-inclusion tasks. In an example of these tasks, children are shown a set of flowers, most of which are roses and the rest of which are another type of flower, and are then asked, "Which is more,

the flowers or the roses?" Young children typically respond that there are more roses, mistakenly comparing subset to subset rather than set to subset as asked. When children answer the question correctly, Inhelder and Piaget (1956) consider them able to use a class-inclusion structure inferentially, reasoning on the basis of the logical relationship that holds between classes and their subclasses. This achievement is considered indicative of concrete operational thought.

Different classification tasks were used by Bruner and his collaborators (e.g., Bruner et al., 1966), as well as by Vygotsky (1962). Nevertheless, their findings parallel Inhelder and Piaget's. For example, in one study Bruner et al. presented 6- to 19-year-old subjects a series of concrete nouns and asked them to tell how each new item in the list was like or different from all the ones they had already heard. Subjects who previously had heard the words *banana* and *peach* next heard *potato*, and were asked how a potato was like a banana and peach and how it was different. Younger subjects favored perceptual solutions, for instance, based on the color of the items. With development, solutions based on function and nominal criteria increasingly predominated. Vygotsky's data on children's ability to classify blocks on the basis of whether they have a common nonsense word on their underside converge on those of Piaget's and Bruner's. Like Inhelder and Piaget, he found that 5- to 8-year-old children reach a point where they can use a consistent criterion. Still, not until they are about 10 years old could they use a classification criterion based on the intersection of two dimensions (e.g., red and square).

We do not dispute that young children's failures on these traditional classification tasks are reliable and well documented. However, by now the same can be said about children's ability to succeed on different classification tasks (for reviews, see Gelman & Baillargeon, 1983; Mandler, Ch. 6, this Volume; Markman, 1989, Ch. 8, this Volume; Rosser, 1991; Wellman & S. Gelman, 1992). Especially pertinent to this chapter are those studies that reveal the potent effect of a knowledge base. When children know a great deal about something, so much so that we can say they have achieved a principled organization of it, there is no question that they can deploy hierarchical classification structures. Our favorite case comes from Gobbo and Chi's (1986) studies of preschool dinosaur experts. These children can sort dinosaurs on the basis of whether they are land-living or not, meat-eating or not, and so on. Since children who know much less about dinosaurs generate data that resembles the results from the traditional classification tasks, it is hard to

escape the conclusion that the key variable here is knowledge, not the presence or absence of classification abilities per se (cf., Carey, 1985). These knowledge effects buttress domain-specific theories of cognitive development.

Problems with the Classical Theory of Concepts.

Logical and empirical challenges to the classical theory of concepts reinforce efforts to find alternative models for our ability to form and use concepts (e.g., Armstrong, Gleitman, & Gleitman, 1983; Carey, 1985; Fodor et al., 1980; Goodman, 1972; Keil, 1989; Markman, 1989; Medin, 1989). The many demonstrations that young children put together items that older children and adults never would, while excluding from their groupings exemplars that are obvious to their elders, can be summarized as follows. Young child cannot correctly "use" concepts when our model of competence derives from the classical definition of what a concept is. According to the classical view of concepts, concepts are defined by their *intension* and their *extension*. The *intension* of a concept is its set of necessary and sufficient fundamental features, none of which can be broken down into smaller sub-features. The *extension* consists of the set of instances that are consistent with the intensional definition, and excludes all instances that are not. Boolean rules of logic organize and generate concepts.

The problems with the classical theory of concepts are well-known. Most importantly, it turns out to be very hard to find examples of everyday concepts that actually have intensional definitions, for example, for which we can articulate the required list of necessary and sufficient nondecomposable features (Armstrong, Gleitman, & Gleitman, 1983; Fodor, 1972). Even baby birds, who cannot fly and have no feathers, are still birds. Second, different people assign the same exemplar to different categories: is a whale a fish or a mammal, and is a tomato a fruit or a vegetable? Third, people treat some exemplars as better examples of a concept than others, as for example rating a robin as a "better" bird than a penguin. Finally, as discussed above, what counts as a relevant feature is concept dependent, and concept salience can vary as a function of context. For example, a snow shovel is more typical in the contexts of Minneapolis and Montreal than it is in the contexts of Los Angeles and Tel Aviv.

The recognition of these problems led to the development of the prototype and family resemblance theories of concepts (Rosch, Mervis, Gray, Boyes-Braem, & Johnson, 1976). The central idea for both is that given a set of

exemplars, we form an abstraction that summarizes them. Novel instances are accepted as further exemplars on the basis of a similarity match to the summary: the more similar an exemplar is to the prototype, the more likely it will be accepted as an instance of the concept. Our ability to categorize faces as familiar is often cited in favor of a prototype or family resemblance theory. Ponder the ability to meet someone you do not know and say, "You must be so-and-so's sister. You sure look a lot like her," despite the fact that the stranger and her sister do not share all of the same features. Consider also our ability to form a concept like *game*, even though we cannot give a classical definition of it that encompasses all cases of games and no cases of nongames (Wittgenstein, 1953).

Despite their improvements over the classical theory of concepts, the similarity-based prototype and family resemblance theories have serious problems as well (see reviews in Armstrong et al., 1983; Fodor & Lepore, 1996; Medin, 1989; Wellman & S. Gelman, Ch. 11, this Volume). For example, the prototype theory fails the test of concept combination. The combination of the representations of a prototypical pet (dog or cat) and a prototypical fish (salmon or trout) does not generate the common belief that the prototypical pet fish is a goldfish. Additionally, there is no a priori definition of similarity and relevance for attributes. The assumption that similarity translates to "close to" fails more often than not. We might say that 3 is more like 4 than 5 because it is closer to 4. But is it still, if we are concerned about the concepts of odd or prime numbers (Armstrong et al., 1983)? Likewise, we might say that a doll is more like a person than is a lobster, because it shares more relevant perceptual features. But this would not be so if we were concerned about the concept of animacy and the capacity for self-generated motion (Massey & Gelman, 1988).

Again, we need an account of why some attributes are treated as relevant for a concept and why others are not. Why do we think that some things are similar to each other even if they are not perceptually alike on their surfaces? What leads us to select a given input as relevant on one occasion but not on another? With core concepts, our answer is: our domain-specific knowledge structures, for both cognitive development and adult concept learning. Indeed, we think some of the best evidence for this position comes from demonstrations of early conceptual competence (see Postulate 4). Our discussion of how domains enter into the definition of similarity begins with a consideration of what count as primary input data for learning.

Characterizing the Learning Environment

Primary Data: Structured, Relational, and Abstract

Within the associationist framework, sensory data are the primary data for concept learning. Sensations that contribute to the build-up of association strength occur frequently, and in proximity to each other in time and space. In turn, associations between existing sensory-based associations are formed; this makes it possible for learners to move on to use perceptual information. Eventually enough examples of perceptual inputs are encountered and learned about to supposedly allow the induction of an abstract concept.

In contrast, our theory views the ability to organize data according to abstract criteria as foundational, not derived. We define primary data in structural terms: a domain's organizing principles lead learners to attend to relations and abstract principles when interacting with the environment, not bits of sensations. This is not to say that sensory data are totally irrelevant and never noticed. It is just that they do not form the foundation for getting concept learning off the ground. Babies use abstract principles about the nature of concepts to find exemplars in the world. Once they have identified exemplars, they can go on to learn whether objects are red, shiny, and so forth (Gelman & Brown, 1986).

Our characterization of primary data for learning in terms of relationships and structures, as opposed to bits of sensations, reverses the traditional developmental rank ordering assigned by associationist theories. That is, the different theories have opposite rankings regarding what kinds of inputs are initially salient and pertinent to both perceptual and cognitive development. Being able to attend to color makes it possible for infants to learn the color of a moving object, but only after the object has been identified as an object at all, distinct from its background (Spelke, 1990). Having nonverbal number-relevant principles makes it possible to attend to the numerosity of a display, no matter what color, shape, size, and so on, the items are.

To review so far: first principles help focus attention on inputs that are relevant for the acquisition of concepts in their domain. It matters not that these first principles are implicit, preverbal, and sketchy in form. What matters is that they are structured. Their active application leads infants and young children to find and store structurally related examples of domain-relevant inputs. Of course, the more examples of data that are present in the child's everyday environments, the greater the opportunity for this kind of learning device to do its work. There is a need for inputs

that share the same expected structure with one another (structural repetition), but there is no requirement for frequent exposure to the exact same inputs (literal repetition). If learners had no mental clues about the surface features of relevant environments, they would have no reason to put together bits of inputs that cohere as objects, that provide clues about causal agents, and so on. Bits of sense experiences do not make an object "out there," let alone an "agent." Selection at this level requires some conceptual organizer. If this is an appropriate way to characterize the nature of relevant data, then we should find that young children and infants are able to respond at abstract levels, and we do (see Postulate 4). Toward the end of the chapter we return to reconcile the conclusion that concepts derive their meaning from a set of related principles with the kinds of facts that lead some to still favor prototype models.

The current literature on mathematics learning adds weight to the theme that learning benefits when the abstract is introduced—alongside the usual concrete examples. A judicious mix of the concrete and the abstract can enhance the development of understanding (Silver, Mamona-Downs, Leung, & Kenney, 1996). Lampert's (1986) successful efforts to bring her 10-year-old students to a principled understanding of multi-digit multiplication algorithms illustrates this point. In one series of lessons she first had fourth grade students (average age = 8.5 years) learn to tell "math" stories, for example, about the total number of children attending parties when 43 children go to 26 parties, versus when 26 children go to 43 parties. Next they were asked to make up stories for multiplication problems. For example:

Teacher: Can anyone give me a story that could go with this multiplication: 12×4 ?

Jessica: There were 12 jars and each had 4 butterflies in it.

Teacher: And if I did this multiplication and found the answer, what would I know about the jars and the butterflies?

Jessica: You'd know that you had that many butterflies altogether.

The class then proceeded to concretely represent Jessica's story. They also took groups apart and made new groups of different sizes, which demonstrated the numerical principles of decomposition and recomposition as well as the distributive and associative laws of multiplication. Eventually, students could make up multiplication algorithms, a clear index of their principled understanding of

multiplication. The children themselves never talked about the distributive law or any other mathematical laws, but they were encouraged to talk in ways that are consistent with the language and structure of mathematics, something that many researchers now consider to be especially important (Brown, Collins, & Duguid, 1989; Gelman, 1991; Silver, Mamona-Downs, Leung, & Kenney, 1996; Stigler & Fernandez, 1995).

Other recent success stories about teaching mathematics share a commitment to introducing the abstract (structural) level of mathematics from the start, mixing it with "concrete" examples (Anderson, Corbett, Koedinger, & Pelletier, 1995). Again, the idea is not to replace experience with doing problems or applications. Rather it is to render these examples mathematically coherent by providing domain-relevant data, including its specialized language. No one thinks that doing this will produce one-trial learning, nor would that be desirable given what we know about learning and retention. But there is good reason to propose that the chances for learning with understanding are increased. The next sections expand on this conclusion.

Learning Structured Information: The Constructivist Learning Laws of Redundancy and Ubiquity

Amongst those theorists in developmental and cognitive psychology who have moved to embrace a structural constructivist theory of mind are those who emphasize the social-cultural aspects of supporting environments, especially the fact that these are organized by members of the community (for recent reviews, see Rogoff, Ch. 14, this Volume, and Rogoff & Chavayah, 1995). Such community and cultural organizations often are cited as reason enough to reject theories such as ours, which give all learners a shared set of innate, skeletal knowledge structures with which to find, interpret, and learn about environments. The thrust of this argument should be familiar by now: there is said to be no need for nativist assumptions because all learners will encounter comparable environments. Again, there are two problems with this counter to the idea that we all share a set of core skeletal domains. We have already discussed one in Postulate 1: the problem of environmental ambiguity and incompleteness. The second problem follows from the fact that individuals interpret environments with reference to what they already know.

Since learners actively participate in their own knowledge acquisition, we have to give up the idea that the frequent presentation of data will suffice for learning to take place. We repeatedly have said that a constructivist theory

of knowledge acquisition is inconsistent with the idea that learning occurs simply because a given set of inputs are frequently present in the real world. This is because we cannot assume that the learner shares with us the same definition of data, let alone whether some datum has been presented X number of times.

In fact, the entire notion of frequency (counts per unit time) is problematic, because it is not clear what counts as a single learning trial. Once we allow learners to contribute to their own cognitive development, we lose the control we may have once thought we had over what counts as relevant. Learners may or may not attend at a given time. A novice's interpretation of the environment "out there" need not overlap with ours. Naive learners might even treat inputs as relevant that we, the "experts," do not think of as relevant. It is possible that even the same input is treated as different on each occasion it is noticed, with different aspects being attended to each time. Most importantly, constructivist learners can themselves make up data that will feed the growth of existing structures. Compelling examples of self-generated data have been provided by those who study the problems of confabulation in eyewitness testimony (Loftus, 1991).

None of these facts fit with an associationist theory of learning that requires frequent exposure to the exact same input. No matter how conceptually competent a child's parents, teachers, and other socializing agents might be, we cannot have a theory of environment that fails to recognize and accommodate the learner's active role in determining what is relevant. To achieve this, we need to develop further the implications of assuming a rational-constructivist theory of mind, one that is founded on the assumption that structures of mind can both define and seek out structured inputs. The structured nature of environments has to be placed alongside the given of a similarly structured mind. The latter is what makes it possible for learners to benefit from the richly structured environments in which they develop.

Given a structural definition of relevance, particular examples can vary in their surface details as long as their structure maps to that of the mental conceptual domain being learned about. For example, $(1+1)$ is equivalent to $(0+2)$, $(90-88)$, and so on. Since we can say that examples within an equivalence class are redundant, we now have a straightforward way to state the laws of input for a constructivist theory of learning: learning is favored under conditions of structural *redundancy* and structural *ubiquity*. The more structurally redundant and ubiquitous examples of domain-relevant information the learner gets, the

more likely it is that the learner's existing mental structures, related attentional inclinations, and potential supporting environment will be mutually compatible.

The principle of ubiquity is *not* the associationist principle of frequency in disguise. The idea is that learners will ordinarily encounter many different exemplars of structurally equivalent inputs. This maximizes the probability that they will recognize enough of them as relevant. The surface details of particular examples typically *differ* from one another, possibly occurring no more than once each. A straightforward count of these details would yield experienced frequency counts of $N = 1$, which is hardly enough to meet the requirement of the associationist law of frequency.

The functions of the ubiquity principle and redundancy principle cannot be separated: they work together as follows. The ubiquitous presence of structurally equivalent (redundant) examples of domain-relevant data is likely to move individuals further along a domain-relevant domain's learning path. This is well illustrated by children's language learning accomplishments. Within a given language environment, children seldom hear the exact same utterance twice. Further, speakers from different cultures vary in the extent to which they simplify their talk to beginning language learners (Schieffelin & Ochs, 1986). Indeed, congenitally deaf children hear neither motherese nor any other variant of spoken language, and whether they see exemplars of sign language varies both within and across countries (Alibali & Goldin-Meadow, 1993; Newport, 1990). Still, both hearing and deaf children all over the world master the syntax of their language group at about the same age, presumably because they encounter many different examples of the surface structure of the syntax that underlies these outputs, and because deaf children somehow find the relevant inputs (Johnson & Newport, 1991).

Language comprehension in general articulates the principles of ubiquity and redundancy. People seldom generate the exact same utterance. Nevertheless, they do produce many acceptable sentences that reflect the operation of their implicit knowledge of syntactic principles. Put differently, sentences are patterns of sound that are isomorphic exemplars of the structure in question. It does not matter whether two examples of relevant data are identical in surface detail, or produced by the same source. They can be offered by peers, adults, television, and so on. What matters is that utterances produced at two different times, by either the same or a different person, reflect a shared structure. When this happens, a child who shares with others the same core domain structure (at least to a considerable

extent) will find the relevant data. To learn their language, children need not hear these data at a given time or from a particular teacher. We hold that the same considerations apply for privileged domains of concept learning, which is why we choose to characterize relevant data in terms of the constraints on the structure of a domain and the principles of ubiquity and redundancy.

The ubiquity and redundancy principles have important practical consequences. They mitigate against the fact that experts' guesses about relevant data sometimes will be wrong, and can even lead to our withholding what young learners would consider to be possible inputs. A theory that allows novices more than one information source in more than one setting makes evolutionary sense, for it acknowledges the possibility of multiple learning routes which are structurally related. This fits well with the important fact that there can be both universals of cognitive development and also culture-specific variability.

Sources of Structured Information in the Learning Environment

We emphasize again that since what novice learners know can differ from what is known by their siblings, parents, tutors, extended family members, and friends, we cannot assume that there are standard countable trials of learning as required by the associationist law of frequency. This shows why it is not a trivial task to develop a theory of the environment for conceptual development that is consistent with any version of a constructivist theory of mind, be it ours, Piaget's, Rogoff's, Vygotsky's, or anyone else's. There are many sources of potentially relevant inputs. Some kinds of learning may lean heavily on one particular source; other kinds will benefit from the potential for redundancy across multiple sources. For example, the idea that one's elders should be respected is surely dependent on inputs from the social-cultural milieu, but probably owes little to the structure of the physical environment. However, the world of physical objects is likely to be the key source for the belief that solid objects cannot pass through one another, even though our perceptions appear to allow this possibility (Leslie, 1988). Individuals might also provide themselves with novel cases of concept data, such as when they contemplate the effect of adding one to a very large number (Hartnett & Gelman, *in press*). Still, from our point of view, any effort to assign more significance to one kind of environmental source than another in the epigenesis of knowledge is akin to asking whether height or width contributes more to a rectangle. With this caveat in mind, we

turn to considering some of the kinds of relevant input that can support cognitive development.

Structure in the Sociocultural Environment

Human beings largely create their own environment. They also tend to structure the learning environment that their children encounter (Rogoff, Ch. 14, this Volume). Sperber (1994) calls this source of experience the "cultural" domain that the learner lives in and interacts with, an extension of his idea of an "actual" domain (see p. 601). The cultural domain consists of the sorts of information that learners experience due to the particular emphases and representational tools of their peers, teachers, and general culture. Sociocultural influences on learning play an important role in the transmission of knowledge, not only in formal educational settings but in informal settings such as parent-child interactions. From our point of view, they are one of the sources of data that children encounter on a given learning path. Greenfield's studies (1993) of how daughters of the Zinacenteco (a Mayan people of Southern Mexico) learn to weave provide a particularly compelling example of this conclusion, in one of the best examples we know of learning that is scaffolded, informal, and relatively error-free.

When Greenfield first visited the Zinacenteco, the mothers were ever-present teachers who continually guided their daughters' efforts to learn to weave. Indeed, they were model scaffolders in the Vygotskian sense, ready to provide explanations and demonstrations that were especially well tuned for their daughters' skill level, so much so as to create the conditions for errorless learning. Some twenty years later, after trucking and tourism became part of the community's ongoing activities, this teaching style has all but disappeared. Mothers are no longer at home—they make fabrics and sell them to tourists—and girls have joined their brothers at the local school. Girls still learn to weave, but not in the same way; much of their learning is self-initiated and self-monitored. Older sisters frequently serve as teachers, but they do not use the scaffolding teaching style. The learner now has to ask explicitly for help from her sister, who often is more interested in other ongoing events. The upshot is that trial-and-error learning has become the rule, not the exception.

Greenfield (1993) expected these changes in teaching style and mode, as well as a change in kinds of fabrics that were made. In order to keep up with tourist demands, women shifted from producing expensive, time-consuming traditional patterns to cheaper and novel patterns. Under

these circumstances, the premium is on novelty and speed. A trial-and-error method of learning is more likely to generate variable outputs, some of which are correct and some of which are not. Since cheaper materials are used, the cost to the group of errors drops considerably. At the same time, girls can be encouraged when they produce interesting variations that appeal to tourists, or when they show a particular talent for weaving.

Lest one treat Greenfield's follow-up study as evidence for the idea that culture is a source of infinite variability, it is important to point out that the girls were still learning to weave, and still producing recognizable patterns. That is, despite cultural changes, there was a shared commitment to the same class of activities and products and this guaranteed that exemplars of the target class of artifacts were available to the learners when they sought them out. Despite dramatic changes in the culture's teaching style, there still were plenty of examples of the structure of the skill to be learned. The cultural unconscious of the Zinacenteco still includes ubiquitous and redundant examples of its commitment to weaving, and the assumption that the skill will be passed from female to female (see also Gelman et al., 1991, on the role of museums as institutions of informal learning).

Thus, when it comes to the arena of cognitive development, it is important to keep in mind the structural universality of some ontological worldviews and social premises (cf., Boyer, 1995; Brown, 1990; Hirschfeld, 1996; Sperber, 1994). For instance, consider that the physical environments in which children are raised vary widely in exposure to animals, for example, a house in the country versus a high-rise apartment in the city. These differences are likely to influence how much children come to know about the core differences between animate and inanimate objects, but nevertheless all children will acquire organized knowledge about these differences. Carey (1995) and Keil (1995) develop a related argument about the cross-cultural evidence for a naive theory of biology.

Structure in the Physical World

Several recent movements in psychology and cognitive science have also called attention to an analysis of the structure of the learning environment. Ecological psychology, founded by J. J. and E. R. Gibson, describes the learning environment as providing an information flow that constantly bombards the perceiving mind with an enormous amount of information, an amount which greatly exceeds the mind's ability to attend to and store

all of it. It argues against associationist theory by showing that the learning environment is structured, and therefore that the senses pick up structured as opposed to punctate bits of information. Three types of structure in the learning environment are particularly important here. First, there are structural regularities in the physical nature of the environment which remain constant over time (Shepard, 1987). Examples include the direction of gravity, the rising and setting of the sun, the fact that solid objects do not pass through one another, and the fact that different surfaces offer different degrees of support. Structural regularities do not mean surface constancies or invariances: for example, the particular time and position of the sunrise is not constant. Nevertheless, there are regularities about the sunrise that can be learned: the fact that the sun does rise every day, and the annual cycle to its north-south variability.

A second type of structure in the environment comes from the division of types of objects into natural kinds. In the case of animate objects, this structure is taxonomic and hierarchical, and has arisen through the multiplication of species during the course of evolution. Animate objects tend to share properties such as having eyes, limbs, and so on. In the case of inanimate objects, diversity has arisen because of differences in raw materials (e.g., granite versus clay) as well as differences in the physical forces that have shaped them (e.g., wind versus water). In both cases, objects belong to structured classes which have many predictable properties and can often be arranged hierarchically (Atran, 1995; Keil, 1995).

A third type of structure in the environment exists at the personal level of the learner: objects in the world interact differently with the learner and offer differing functional possibilities. Animate objects tend to move and interact with their environment, raising the possibility of threat or benefit to the learner. Inanimate objects can also move, but their motions result only from external forces rather than from internal motives. The Gibsons' concept of affordance is closely related to the functional aspect of this type of structure: only certain objects in the environment can be walked on, picked up, or thrown. Children quickly catch on to the relevance of these kinds of affordances. For example, 3- and 4-year-old children say they cannot touch the sun because it is too far away, and if it were not, it would be too hot anyway (Gelman, Spelke, & Meck, 1983).

Again, we must emphasize that the types of structure in the learning environment just described need not be concrete and directly perceivable in the empiricist sense.

Some regularities are temporal, such as the seasons of the year. Other regularities are internal, such as concepts of causation and mechanism which underlie reasoning about animate versus inanimate motion. Still other aspects of structure are abstract, as in the case of numerical regularities. None of these examples of structure could be learned by traditional empiricist means, in which punctate sensations have to be proximate to each other in time and space in order to be associated.

Structure in the Mind

There are a variety of ways in which the mind is capable of generating data for its own benefit, a capacity mentioned several times above and which Piaget referred to as "logical experience." Self-generated data certainly play an important role in language learning: throughout infancy, babies babble to themselves as they try out and refine new speech sounds. A similar role is played by early stages of birdsong learning (see discussion in Postulate 6 above of Marler's work in this area). Gelman and Hartnett (in press) provide a variety of examples of how young children use thought experiments to generate counting-relevant data: counting to some large number, or continually adding one to a number they already think is rather big. Indeed, it was parents' reports of their preschoolers wondering about whether the numbers ended that led Gelman and her students to study children's learning about the successor principle for natural numbers.

We are only beginning to learn about this self-generated source of cognitive development, especially the conditions that engender the spontaneous generation of questions and inferences that can attract appropriate inputs from those who are more knowledgeable. Brown and Campione's (1996) "communities of learners" in elementary schools, like microgenetic studies of a focused learning problem (e.g., Karmiloff-Smith & Inhelder, 1974-1975; Kuhn et al., 1995; Siegler & Crowley, 1991) are extremely promising on this front. They provide detailed accounts of acquisition curves as well as the kinds of input learners both generate and use.

Mental Learning Tools: Beyond Initial Enabling Constraints and Core Domains

Up to this point our focus has been on the role of enabling constraints and the nature of conceptual domains. In the remainder of the chapter we turn to what we call "mental learning tools." By mental learning tools, we mean an

armament of learning mechanisms such as structure mapping, the computation of frequencies and contingencies, imitation, template tuning, parameter setting, pattern seeking, action-planning, and so on.² For us, mental learning tools contribute to the active construction of knowledge, and are neither domain-specific nor domain-general in the usual sense. They are more specific than truly domain-general processes such as formal logic, and yet more general than the enabling constraints in core domains.

Mental learning tools are more "central" than enabling constraints in Fodor's (1983) sense of peripheral versus central modularity. This is because they often take as inputs the data which has already been selected as relevant by domain-specific enabling constraints. A learning tool is applied whenever the learner is faced with data having the particular structure with which that learning tool "resonates" (Shepard, 1994), in something like what Keil (1994) refers to as a "mode of construal" and Dennett (1987) refers to as a "stance." For instance, if a task requires mastering the frequency with which particular exemplars of a concept appear in the learner's environment, then the frequency computation device will be engaged (see below).

Mental learning tools are therefore not strictly domain-specific, in that they can operate on structured information from more than one domain, but neither are they completely domain-general, since they resonate only with a specific type of informational structure. They provide a way to construct higher-order combinations from the outputs of different domain-specific learning mechanisms, in order to produce more flexible and general learning. It is combinations of this sort which we believe create the *appearance* of domain-generality. Higher-order combinations of information by learning tools also address the issue of evolvability raised by Tooby and Cosmides (1992) in their sweeping rejection of domain-general processes. They argue that truly domain-general solutions to adaptive problems are always weaker than problem-specific solutions, and therefore could not have evolved. Tooby and Cosmides' objection does not apply to learning tools, however, since

they are not truly domain general. Learning tools create potentially useful extensions of domain-specific processes. This potential usefulness creates a gradient which satisfies the conditions of evolvability: some benefit is better than none, and therefore gradual improvements in learning tools could evolve (see Dawkins, 1995, Ch. 3).

Our list of mental learning tools above might strike some as unusual, even perverse. It brings together processes that are sometimes pitted against each other, for example, imitation and parameter setting. Discussions of imitative learning as a domain-general process have often been embedded in noncognitive social learning theories (e.g., Bandura & Walters, 1965; Miller & Dollard, 1941). This fact surely contributed to treatments of imitation in parameter-setting theories as a poor stepsister of general associative learning mechanisms (Chomsky, 1959, 1965; Piattelli-Palmarini, 1989). Even so, what might be true about syntax acquisition need not apply for all domains of learning. We agree with Premack and Premack's (1995) thesis that imitation serves young human learners' efforts to accomplish many knowledge acquisition tasks, including norms of human interaction. In doing so, we reject the idea that imitation is best accounted for within a content-neutral S-R social learning theory like Bandura and Walters'. We share with Piaget (1929) and other cognitive developmentalists the position that imitation serves learning. We consider it an extremely potent example of structure-mapping, the first of two mental learning tools described in the next section. Frequency computing is the second learning tool featured.

Structure Mapping

The ability to map from existing mental structures to new structures during knowledge acquisition is among the most important of the mental learning tools. This follows from our assumption that principles of a domain make possible the assimilation of domain-relevant data, that is, data that share a common structure with the domain's principles. In core domains, the skeletal principles themselves define initial representations which become the repository of all data whose structure can be mapped to them. Structure mapping serves learners' abilities to identify relevant inputs when going beyond old learnings, given the ever-present tendency to apply existing mental structures to new learnings. A similar mechanism can be invoked to address Wilson and Sperber's (1986) account of relevance.

Our claim that structure mapping is a key learning tool has an important implication: learning will be a function of

² We are much aware of the ability of humans to develop and use a host of metacognitive learning skills and artifacts with varying degrees of facility, including physical, symbolic, computational, and communication tools. Our focus here is not on these metacognitive tools that learners can deploy to their benefit once they are mastered, but rather on online mental learning tools whose use does not depend on their intentional deployment.

the degree to which existing mental structures (be they in core or noncore domains) overlap with the structure of the input. Environments that share relevant structural relations with existing domains of knowledge are ones that are conducive to rapid knowledge acquisition. Our premise is that learners have a better chance of assimilating an input if there is potential for a structural map between what is to be learned and what is already known. The fit can also be dependent on whether there is a match between the number of common *structural units* a student can process and the number embodied in the data (cf., Case, 1992; Halford, 1993). In the absence of such potential structural maps and fits, there is a risk that the input will be ignored or mistakenly assimilated into the existing (inconsistent) structure and knowledge base (Gelman, 1991).

Work on analogical reasoning provides especially compelling examples of the foregoing ideas about structure mapping. There is an emerging consensus that successful analogical transfer depends on the extent to which structure mapping is possible. The probability that transfer will occur is very high if learners achieve representations that are structural isomorphs but very low if learners have to rely on surface cues of perceptual similarity (Brown, 1990; Gentner, Rattermann, & Forbus, 1993; Holyoak & Thagard, 1995). For instance, by taking advantage of very young children's principled knowledge about causality, Goswami and Brown (1990) were able to illustrate that children solve analogies of the $a:b::c:d$ form when a common causal transformation forms the basis of the analogy. In another example of analogical reasoning and transfer, Catrambone and Holyoak (1989) presented subjects with multiple analogous word problems and measured transfer to the solution of a superficially different word problem; transfer was facilitated when the problems were worded so as to emphasize their structural similarities.

The potential role of a structure-mapping learning tool is also featured in the literatures on mathematics and science learning. There are a number of differences between mathematics classes in Japan and the United States for 6- and 7-year-olds, including how children are taught number facts such as multiplication tables (Stevenson & Stigler, 1992). In the United States, children often are introduced to their number facts with manipulable objects so as to render the symbolic level "concrete." The roots of this practice lie in the mistaken belief that young children are limited to concrete (non-abstract) reasoning. Although the same manipulable materials can be found in Japanese classrooms, they are used in a very different way; Japanese

teachers relate them to abstract mathematical representations in their discussions about the many possible ways to state particular number facts. Both errors and correct answers are included in the discussion, in an effort to encourage all children to think of different but structurally equivalent ways to state a single mathematical problem, for example, $2 + 5 = 7$ and $1 + 1 + 1 + 1 + 1 + 1 + 1 = 7$. Although the left-hand sides of these two equations are perceptually very different, they are structurally equivalent within the domain of natural number arithmetic, making structure mapping to the principles of the domain possible, and thereby facilitating a deeper understanding of number facts beyond mere rote memorization. Efforts to promote structure mapping and overall coherence are also characteristic of Japanese lesson plans, and are a leading hypothesis proposed by Stigler and his colleagues (e.g., Stevenson, Lee, & Stigler, 1986; Stigler & Fernandez, 1995) for why Japanese elementary school math lessons lead children to better understandings than American lessons.

Gallistel and Gelman's (1992) conclusion that preverbal counting processes are isomorphic to verbal counting processes serves as another example of how the domain of number can take advantage of structure mapping. Preverbal principles provide a framework that makes the verbal counting process intelligible to the young learner; a structure mapping learning tool makes possible the assimilation of the count list encountered in the child's environment (Gelman, 1993). Gelman and Gattis' (1995) review of mathematics learning contains other examples of structure mapping as a learning tool. These include Lampert's (1986) successful efforts to bring her elementary-school students to a principled understanding of multidigit multiplication algorithms (see p. 607), and Nesher and Sukenik's (1991) successful program for teaching high school students about ratios and proportions.

Further examples of the workings of structure mapping as a learning and problem-solving tool are found in the literature on science learning. White's (1993) ThinkerTools program for teaching Newtonian mechanics to sixth graders takes as given the assumption that science learning should be built around causal principles and relations. Studies from the literature on expertise and novice-expert differences also illustrate how a domain can take advantage of a structure-mapping tool. Chi, Feltovich, and Glaser (1981) asked novices and experts to sort physics textbook problems in any way they wished. Novices did so on the basis of the perceptual aspects of the diagrams or the apparatus, for example, inclined plane, balance beam, and so on.

Experts instead classified the problems on the basis of the underlying physics principles needed to solve the problem, for example, Newton's second law. Put differently, experts were able to map the organizing principles of their (non-core) domain of knowledge to the different problems. Since the novices did not share such principled knowledge, they could not achieve a structural map based on the laws of physics. This being so, they could either use a perceptual default strategy to classify problems on the basis of common surface information, or map the input to whatever existing mental structures they had available. Both kinds of solutions occurred: the former were classified as examples of perceptual solutions, the latter as examples of misconceptions.

Young children have a ubiquitous tendency to persist until they get something right (Brown et al., 1983; Gelman & Brown, 1986). This means that in addition to their tendencies to apply their existing mental structures, they have a way to monitor the relationship between their outputs and the target they aim for. Our designation of structure mapping as a learning tool provides an important piece of the solution to the problem of accounting for self-initiated and self-guided learning within a domain. A structure-mapping device can provide at least a *match/no-match* test of whether a solution is correct. Repeated efforts can be scrutinized in the same way, and accepted or rejected to the degree that they map structurally to the principles of the domain governing the learning or problem solving. Learnings that can take advantage of existing structures are at an advantage in this account, a prediction that is supported by the wealth of data showing that learning is always better when there is an existing mental structure.

The foregoing leads to a principled statement about structure mapping: Learning is a function of the degree to which existing knowledge structures can be mapped to—in Piagetian terms, projected onto—the to-be-learned materials. Theory and ongoing research on the relation between early domain knowledge and later learning lends strong support to this conclusion (Hartnett & Gelman, *in press*).

Learners' proclivities to map available structures to environments helps explain some rather precocious abilities in young children. Before they are taught to write, they can generate different plans of action in response to a request to either write a word or draw a picture for a given line drawing (Brenneman, Massey, Machado, & Gelman, 1996). They can also develop writing systems (Tolchinsky-Landsmann, 1990), and they can distinguish between a string of marks on paper that are "good for writing" as opposed to "good for numbers" (e.g., Lee & Karmiloff-Smith, 1996).

How do the young even begin to sort out the fact that there are different kinds of marks on paper? We suggest that young children's structure mapping tendencies serve them well in this case. The idea is that each of the symbol systems has its own structure and related constraints. For example, we know that young children have implicit knowledge about the structure of inanimate objects and language. The former have bounded surfaces and are solid. Drawings of objects map these characteristics, at least often enough. An orange is drawn as a continuous circle and filled in with the color orange. Language is represented as a sequence of sound units; print (at least in the cultures studied) consists of a sequence of marks with spaces between them, and so on. Such differences in structural relations do not begin to define the full range of our implicit knowledge about objects and speech, on the one hand, and drawing and writing conventions on the other hand. Still, they might suffice from the viewpoint of young learners whose goal is likely to be limited to an attempt to distinguish between the kinds of marks on paper that they encounter. This goal can be served by their omnipresent tendency to engage in structural mapping. Young children's beginning representations of the difference between drawing and writing are hardly complete. What matters is that they exist at all. Once they do, a learning path opens up—and in this case there will be many eager to encourage and support movement along it (Cole, 1996).

Frequency Computing

As indicated, skeletal principles in core domains draw attention to the class of relevant inputs, and organize the assimilation and early representation of noticed cases. Skeletal structures start to accumulate flesh as structured examples are assimilated. In addition, they take advantage of an automatic (nonconscious) ability to keep a running frequency count of encountered exemplars and their relevant aspects. Such a learning tool contributes to the build up of knowledge of the predictive validity of the different attributes of encountered exemplars. For example, certain surface properties and form attributes characterize animate objects, as opposed to different properties and attributes which characterize inanimate objects.

The registration of attribute frequencies and the computation of their predictiveness, that is, of the contingency between a given concept and the possession of a given surface attribute, is carried out by what we call a frequency/contingency learning tool. This learning tool is a good example of the middle ground described above between domain generality and domain specificity; it

operates specifically on frequency data, but it also performs the same frequency-computing function in many different domains.

There is good evidence that animals and humans of all ages keep track *automatically* of the frequency of relevant events and objects (Gallistel, 1990; Hasher & Zacks, 1979; Marcus et al., 1992). For example, Hasher and Zacks (1979) showed children ages 5 to 8 years a series of pictures, in which each picture appeared 0 to 4 times. Afterward, children in all age groups were highly and equally successful at reporting how many times a picture had been shown, despite not receiving any instructions to keep track of this information. Similarly robust abilities to pick up frequency information about objects or events abound. Hasher and Zacks (1984) have documented frequency learning across populations (college students, learning-disabled children, and depressed and elderly persons), as well as across a wide range of variable-frequency materials (letters of the alphabet, familiar words, surnames, and professions). Marcus et al. document people's ability to keep track of the different frequencies of irregular past tense and plural words. Infants' abilities to adjust the frequency with which they suck or turn their heads in order to achieve presentations of sounds, well-focused photographs, mobiles they control, and so on, add to the list of cases where we find the mind keeping track of frequencies of items or events of interest (Watson, 1987).

Saffran, Aslin, and Newport (1996) provide another impressive demonstration of frequency detection and learning. They presented 8-month-old infants with just two minutes of synthesized speech in a monotone female voice. The speech stream contained four three-syllable "words," repeated in random order with no pauses. Despite the lack of pauses, the infants were able to detect the boundaries between words in the continuous speech stream. The only information available to them for this task was frequency information: the transitional probabilities between particular pairs of syllables (higher within words than across word boundaries). In one experiment, infants reliably distinguished words from nonwords where the nonwords consisted of familiar syllables in an unfamiliar order. In a second experiment, they successfully made an even more difficult distinction, in which the nonwords consisted of the final syllable from one word and the first two syllables from another word. The syllable order of each nonword was therefore familiar, but statistically its syllables did not correspond to a word. Saffran et al. concluded that these results were attributable to an innately biased statistical learning mechanism. This mechanism could be language-

specific, or (as we believe) more generally applicable to distributional analyses of environmental stimuli.

We are sure that readers know which is more frequent, white or green cars and are very impressed with Macario's (1991) preschool children who could play his "what will we take on a picnic" game—they were able to generate possible foods on the basis of a color cue. We endorse the thesis put forth by Hasher and Zacks (1979, 1984), that learning about the frequency of *noticed* objects and events occurs automatically and without awareness. So apparently do Tversky and Kahneman (1973) who simply assume that we can use base rate information. In fact, we believe that humans share this extremely potent learning tool with other animal species. Gallistel (1990) reviews evidence that animals in classical and instrumental conditioning paradigms are learning the rate/frequency of reinforcement and its contingency on available cues, rather than the associative pairings predicted by associationist theory. From raw frequencies the animals are computing contingencies, that is, the extent to which the frequency of reinforcement in the presence of a conditioned stimulus is different from the frequency observed in its absence. As with humans, these computations are automatic and continuous.

Once again our use of a term, in this case, frequency, admits possible misinterpretations. In our view, the foundational structure in core domains always comes from the skeletal principles embodied in enabling constraints; frequency data about relevant encounters are *subsequently* recorded and attached to that existing framework. Frequency information does not help the learner to recognize encounters in the first place. Thus, our idea that a frequency counting computational device is a mental learning tool is not a variant of the associationist law of frequency governing the learning of associations. Within association theory, frequency serves to build associative strength; it is not specifically encoded and represented, nor does it feed a device that keeps a running total of frequency per se. Certain associations are stronger than others because they have had the benefit of more frequent encounters with particular pairings of stimuli, longer rewards, and/or rewards at shorter delays. That is, many different factors combine to determine associative strength, but the factors contributing to that strength, for example, frequency, are not represented by associative strength and therefore are not recoverable as inputs for learning.

Thus, to say a cue has predictive validity is not to say that it is defining; lettuce does not have to be green, and a green leaf does not guarantee the presence of lettuce. Similarly, the distinction between a malleable versus a rigid form is

strongly correlated with the animate-inanimate distinction. To be sure, the rigidity cue has considerable predictive validity for animacy/inanimacy, as do attributes such as uniform versus variable surface textures, and the presence or absence of limbs, eyes, ears, and so on. Although none of these are defining it still helps to learn their relative frequencies and related contingencies. Such computations allow us to make an informed guess about the animacy status of a novel or unidentified item, and then to check if the guess is consistent with the requirements of the domain. Informed guesses can be disconfirmed by subsequent information which violates core principles about a domain; no matter how much something looks like a rock, we will no longer believe it to be a rock if it gets up and walks away. And no matter how unlikely a particular example of a category might be, we can accept it if it can be assimilated to the domain's principles. Thus, a green lemon is still a lemon, and a three-legged dog is still a dog.

Parenthetically, the above makes it possible to make sense of the data that are used to favor a prototype theory of concepts. These are the data that show learners have knowledge of particular relevant features as well as their centrality (see Schwartz & Reisberg, 1991, for an excellent review). For example, wings are particularly relevant features of birds, and robins are more "central" exemplars of birds than are ostriches. From our point of view, these studies provide evidence that we do have a frequency-counting mental learning tool. This is why high-frequency features are more memorable, and why exemplars with many high-frequency features are judged better examples of a category. How such frequency information is used depends on the domain in question. Individuals will not say that a high-frequency irregular verb is a better example of verb than a relatively novel verb with a regular past tense (Marcus, 1996).

The proposal that a mental learning tool computes the frequency of relevant encounters converges with conclusions drawn by other authors. Schwartz and Reisberg (1991) suggest that we may need a three-part theory of concepts, in which "concepts are represented by a prototype, some set of specifically remembered cases, and some further abstract information" (p. 391), where the parts all interact to accomplish correct similarity judgments and inferences. In our account, the recorded knowledge of frequencies and contingencies underlies subjects' abilities to answer questions in ways that make them look like they learn prototypes and some salient domain-relevant exemplars. Keil (1995) has proposed that "concepts in theories"

structures are supplemented by domain-general feature tabulation processes. Armstrong, Gleitman, and Gleitman (1983) concluded that we know the difference between saying an object is an instance of a concept, versus characterizing it as a good or bad instance. More generally, our account provides a way to reconcile these response patterns with the compelling arguments against the idea that concepts are based on prototypes (Fodor & Lepore, 1996).

Further converging lines of thought exist with respect to children's understanding of causality. Bullock, Gelman, and Baillargeon (1982) argue that causal principles lead children to search for causal mechanisms and assimilate causally relevant information about events, including the cue value of spatial and temporal cues. Ahn, Kalish, Medin, and S. Gelman (1995) conclude that information about covariation and about causal mechanisms play complementary roles in our decisions about causes. Cheng (in press) shows that people relate their computations of contingency to their beliefs in causal principles. Finally, several authors have offered accounts of how children learn to classify moving objects as animate or inanimate based on the causal conditions of animate versus inanimate motion (Gelman, 1990; Gelman, Durgin, & Kaufman, 1995; S. Gelman, & Gottfried, 1996; Williams & Gelman, 1995).

Mental Learning Tools and Later Learning: Structure Mapping Can Help or Hinder

A central tenet of a structural constructivist theory of mind is that existing mental structures influence later learning, whether they be the result of prior learning or of enabling constraints in core domains. Learning is fastest in core domains, when enabling constraints are engaged by the types of inputs they are functionally designed to "expect," allowing the rapid uptake and interpretation of information. Learning is slowest when no existing structures are close enough to the structure of the new information to be useful (e.g., tax law, calculus). In such cases, the learner has the double burden of learning both the information and the structure with which to interpret and store it. As the literature on expertise acquisition in noncore domains has shown, novice learners often settle into nonprincipled organizing structures which appear to reflect fall-back strategies such as perceptual similarity (e.g., Gibson & Gibson, 1955); with enough experience and concentrated study, learners are able to reorganize their structures in a principle-based manner which reflects the underlying regularities in the knowledge domain. For instance, as noted earlier, novice physics students appear to structure their

knowledge of physics around formulas and perceptual similarities, such as the presence of pendulums or inclined planes. In contrast, expert physicists structure their knowledge around underlying physical principles (Hardiman, Dufresne, & Mestre, 1989; Larkin, McDermott, Simon, & Simon, 1980).

Between these two extremes of particularly fast and particularly slow learning lie instances in which some partial mapping is possible between existing knowledge structures and the structure of the information being learned. When this mapping is particularly accurate and consistent, analogical reasoning provides an important mental tool for rapid uptake of the new structure (Gentner, 1989; Holyoak & Thagard, 1995; Vosniadou & Ortony, 1989). There are cases however where existing knowledge structures actually hinder the acquisition of new knowledge. This can occur when entities and concepts which have previously been learned about within one structure, particularly as part of a core domain, are being learned about within a new structure (e.g., Slotta & Chi, 1996). When entities and concepts are shared in the new structure, the learner is likely to assume that the old structural relations between them are still valid. If the new structure turns out to be inconsistent with the old structure, the learner is in for a hard time. We illustrate these implications with an example of our effort to document the contrast between easy and difficult mappings.

An Example: Why Learning about Fractions Is Difficult and Infinity Is Not

Existing knowledge structures may or may not facilitate the learning of new knowledge structures. When the structure of the to-be-learned data does not map to the structure of what is known, existing knowledge can become a barrier or obstacle to new learning. That is, when there is inconsistency between the intended interpretations of inputs for new learning and the actual way a learner is likely to interpret the data, new learnings may not occur. For better or for worse, our constructivist minds sometimes can put new learning at risk, not because we intentionally go astray or intentionally misinterpret inputs, but rather because we cannot help but interpret inputs in terms of what we know and/or believe—just as in the Müller-Lyer illusion, when we see one stick as longer than the other when in fact they are equal. We find consistency even in the face of inconsistency. We do not do this to be ornery, any more than we intentionally err in length judgments when shown the Müller-Lyer display. In both cases, the interpretation of

data follows from structures of the mind which continually attempt to make sense of the world. As we have seen in various sections of this chapter, such structures usually serve us well, helping us select and organize relevant inputs, making sense of the novel, and so on. That they can also lead us astray when we encounter certain input conditions is to be expected. This is one of the pitfalls of the fact that our minds do have knowledge structures.

The idea that there will be some structure-environment pairings that lead to “trouble” is not unique to constructivist theories of mind. Armies have learned to have soldiers break stride when crossing bridges in order to minimize the likelihood that the bridge will collapse. Bridges collapse when they resonate at their natural frequency and it is possible to generate this frequency with certain marching paces. The point for developmental psychologists—indeed all psychologists who care about knowledge acquisition—is this: Even when we lack the structures needed for the correct interpretation of given inputs, we interpret them nevertheless. At such times, the fundamental tendency of our constructivist minds is to apply existing structures, be these about knowledge domains, beliefs, social schemes, etc. Under such conditions, our minds can therefore lead us astray.

There is abundant evidence that schooled and non-schooled individuals all over the world share a common theory about numbers, given that there are some number-relevant activities within a culture. These can include tailoring, cab driving, selling, shopping in supermarkets, or going to Weight Watchers, to name a few. The crux of this theory is that “numbers are what you get when you count.”³ Children in the United States start to receive formal instruction about fractions in kindergarten or Grade 1. Number lines are a ubiquitous feature of kindergarten classes. Blackboards prominently display fraction numerographs and numerlogs, for example, $\frac{1}{2}$ and “one half.” Textbooks include illustrated lessons on the labels of X equal parts of shapes, where X is at least 2, 3, and 4 and the related fractions are $\frac{1}{2}$, $\frac{1}{3}$, and $\frac{1}{4}$. They also offer instruction on how to label the X equal parts, for example, “this circle has 2 equal parts, each is one half of a circle,” “We can write *one-half* as $\frac{1}{2}$,” and so on. Finally, children have some opportunities to mix fractions with whole numbers, in either fractional

³ There are disagreements about the developmental origins of this theory, but not whether it is shared by the time children start school.

notation or natural language (e.g., $1\frac{1}{2}$ and "one and one half").

In contrast, young children are almost never taught about the successor principle for natural numbers, that is, that every natural number has a successor and therefore that there is no last number. However, the mathematical structure of the natural number successor principle is consistent with what children already know about numbers, whereas the mathematical structure for the notion of a fraction is not. For these reasons, we (Gelman, Cohen, & Hartnett 1989; Hartnett & Gelman, in press) reasoned that despite these different schooling opportunities, young children would (a) come to understand the successor principle with relative ease; and (b) make systematic interpretative errors about fractions.

Mathematically, a fraction is defined as one cardinal number divided by another. Since division is an operation that is outside the knowledge of the young child, the idea of a fraction as a number is not conceptually consistent with the counting principles. The definition maps to a tripartite symbol, for example, $\frac{1}{2}$ is made up of symbols for 1, 2, and division. Although the rule of one-to-one correspondence of the counting principles readily maps to single entities like digits, nothing in the counting principles can be mapped to the symbol for a fraction, be it in fraction or decimal format. To make matters even worse, there is no successor principle for fractions; that is, there is no "next" fraction. This follows from the fact that there are infinitely many numbers between any two rational numbers.

The expected differential successes with successor and fraction tasks have indeed been found (e.g., Gelman, 1991; Gelman, Cohen, & Hartnett, 1989; Hartnett & Gelman, in press). As expected, children between 5 and 8 years of age quickly catch onto the fact that they can keep adding 1 to whatever number they are thinking about. In contrast, children within the same age range consistently misinterpreted inputs meant to represent fractions. For example, when Gelman, Cohen, and Hartnett (1989) asked kindergarten, first-, and second-grade children to place circle representations of fractions on a special number line that represented whole numbers with sets of N circles, the majority of the children used whole number count strategies to solve the task. When asked to put a card with $1\frac{1}{2}$ circles on the number line, most children placed it at the position for 2 since the card had two objects on it. Similarly, many children placed all of the representations of unit fractions (one-fourth, one-third, one-half) at the position for 1 on the number line, often telling us that each was "one thing."

Further, although these same children pointed correctly to the card that "showed one half," when their choice was between $\frac{1}{2}$ and $\frac{1}{4}$, they systematically misordered these same two fraction symbols, choosing $\frac{1}{4}$ as larger. Similarly, they seldom read $\frac{1}{2}$ correctly as "one half." Instead they came up with a variety of alternatives, including "one and two," "one and a half," "one plus two," "twelve," and "three." The most frequent misreading was "one and a half." These are but some examples of early conceptual difficulties with fractions and concepts that are related to the underlying mathematical structure.

There is an ever-growing literature that adds weight to our proposal that children's knowledge of natural numbers (a core domain) serves as a conceptual barrier to later learning about other numbers and their mathematical structures, for example, fractions, ratios, proportions, multiplication, and division (see, for example, Leinhardt, Putnam, & Harttrup, 1992; Nunes & Bryant, 1996; Sophian, 1994). Many children eventually do achieve a mathematical understanding of fractions despite the difficulty. Gelman (1991) suggests that they accomplish this by finding a series of local structural commensurates between their knowledge of the counting numbers and bits of knowledge about fractions. In the absence of an existing relevant conceptual structure, these local mappings serve as mental stepping stones which can begin to assimilate relevant data, and thus for some learners to begin to fashion a skeletal structure for learning. But since the learning task requires the simultaneous acquisition of both the relevant structure and the entities that give the structure its content, it should be no surprise that it takes a long time.

CONCLUSION

We have argued for a theory of concept learning and cognitive development that combines converging lines of theoretical and empirical work from across the cognitive sciences. These influences include deep empirical and theoretical challenges to previous theories of learning and cognition; the recognition of domain-specific and modular learning mechanisms in humans and animals; and an appreciation of evolutionary perspectives on learning. Because so many terms are shared across very different theories of learning, we have emphasized the theory-dependent nature of terms, particularly "constraint." Constraints of one form or another are embedded in all learning theories, not just those which endow learners with innate mental structures. We

believe that among the alternatives, a view of constraints as learning enablers is best suited to charting the complex interaction between biological and environmental influences on development. Distinguishing between core and noncore domains facilitates a more productive approach to a variety of long-standing controversies: the degree of innateness in mental architecture, the existence of cognitive universals, and the helping/hindering relationship between early and later learning.

In his review of evolutionary psychology's attempts to serve as a new grand theory for the social sciences, Premack (1996) expressed the opinion that "the cutting edge of [developmental psychology] is entirely modular . . . the villains have already been met and, as much as villains can ever be, were slain." While this may be true of traditional but now empirically untenable learning theories, our impression is that many underlying tenets of Empiricism and associationist theory have survived the demise of behaviorism. Domain-general, equipotential, anti-abstract, and extreme anti-nativist sentiments are alive and well in cognitive science, particularly within many information-processing and connectionist accounts of learning. For those within these camps, therefore, our theory of concept learning still represents a difficult theory change, containing many counterintuitive aspects; all we can propose is that intuition is often a poor guide when it comes to understanding both adaptations (Dawkins, 1987) and psychological theories (Cosmides & Tooby, 1994b). This is particularly true of theories whose support derives even in part from evolutionary arguments, and which therefore commonly face what Dawkins (1987, p. 38) describes as the "Argument from Personal Incredulity." Yet unless we wish to return to Descartes' dualism, we must avoid the inconsistency of accepting adaptive evolutionary arguments for anatomical structures but rejecting them for mental structures. The brain with all its functional complexity and differentiation is a product of natural selection just like the rest of us.

So does our rational-constructivist account constitute a new grand theory of learning? We do not think so. Our attention to terminological definitions and criticisms of traditional views should not be taken as a shield which allows any criticism to be deflected as "misunderstandings." We are just beginning to draw many complementary threads together, to recognize their implications, and to attempt to weave a consistent theory of learning which fits the available evidence better than its predecessors. Our contribution is not limited to the proposal that innate

skeletal constraints enable learning in core domains: the theory itself can be seen as a skeletal structure to which we hope future accounts of learning will cohere. Much work remains to be done in order to develop the idea of mental learning tools, and our hope is that such work will prove fruitful in exploring the interactions between domain-specific and domain-general learning mechanisms. Further development of the theory must continue to rely on converging lines of evidence, based on data from infants, different cultures, animals, atypical development functions, systematic cross-task manipulations, micro-genetic analyses, and theoretical predictions about the nature of variability and the role of learning environments. It also remains to be seen to what extent explanations of additional types of learning beyond concept acquisition can be accommodated by the framework: declarative fact learning, event memory, perceptual learning, implicit learning, motor skill learning, and so on.

In closing, we call attention to the scientific goal of parsimony. Learning theorists' traditional interpretation of parsimony, borrowed from physics, has always been to postulate a small set of simple and domain-general learning mechanisms which are minimally constrained by a few domain-specific preferences. We believe that the theoretical and empirical considerations which motivate our effort call for this definition of parsimony to be turned upside down. We view the mind as a collection of complex, functional, domain-specific adaptations tied together by higher-level adaptations including partially domain-general learning tools. Constraints enable learning, rather than limiting a supposedly general-purpose learning mind. Conceptual learning in core domains would be impossible were it not for the presence of domain-specific enabling constraints which solve the deep problems of selectivity and ambiguity. This view may be counterintuitive for some, but it fits well with the available evidence and with an evolutionary perspective on the origins of the mind. In this new view of parsimony, domain-general learning processes should be seen as the exception rather than the rule.

ACKNOWLEDGMENTS

The final draft of this chapter was written at New York University where Rochel Gelman was a Visiting Scholar in the Psychology Department. We thank Stephanie Reich, Randy Gallistel, and our editors, Deanna Kuhn and Robert Siegler, for their insightful comments on earlier drafts.

REFERENCES

- Ahn, W., Kalish, C. W., Medin, D. L., & Gelman, S. A. (1995). The role of covariation versus mechanism information in causal attribution. *Cognition*, 54, 299-352.
- Alibali, M. W., & Goldin-Meadow, S. (1993). Gesture speech mismatch and mechanisms of learning: What the hands reveal about a child's state of mind. *Cognitive Psychology*, 25, 468-523.
- Anderson, J. R. (1995). *Learning and memory: An integrated approach*. New York: Wiley.
- Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. *The Journal of the Learning Sciences*, 4, 167-207.
- Andrews, F. E. (1977). *Numbers, please*. New York: Teacher's College Press, Columbia University.
- Armstrong, S. L., Gleitman, L. R., & Gleitman, H. (1983). What some concepts might not be. *Cognition*, 13, 263-308.
- Aronson, L. R. (1951). Further studies on orientation and jumping behavior in the gobiid fish, *Bathygobiis soporator*. *Annals of the New York Academy of Sciences*, 188, 378-392.
- Atran, S. (1994). Core domains versus scientific theories: Evidence from systematics and Itza-Maya folkbiology. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 316-340). New York: Cambridge University Press.
- Atran, S. (1995). Causal constraints on categories and categorical constraints on biological reasoning across cultures. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 205-233). New York: Clarendon Press/Oxford University Press.
- Au, T. Kit-Fong, & Romo, L. R. (1996). Building a coherent conception of HIV transmission: A new approach to AIDS education. In D. L. Medin (Ed.), *The psychology of learning: Advances in research and theory* (pp. 193-241). New York: Academic Press.
- Au, T. Kit-Fong, Sidle, A. L., & Rollins, K. B. (1993). Developing an intuitive understanding of conservation and contamination: Invisible particles as a plausible mechanism. *Developmental Psychology*, 29, 286-299.
- Bahrack, L. E., Walker, S., & Neisser, U. (1981). Selective looking by infants. *Cognitive Psychology*, 13, 377-390.
- Baillargeon, R. (1986). Representing the existence and the location of hidden objects: Object permanence in 6- and 8-month-old infants. *Cognition*, 23, 21-41.
- Baillargeon, R. (1995). Physical reasoning in infancy. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 181-204). Cambridge, MA: MIT Press.
- Baillargeon, R., Kotovsky, L., & Needham, A. (1995). The acquisition of physical knowledge in infancy. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary approach* (pp. 79-116). Oxford, England: Oxford/Clarendon Press.
- Baillargeon, R., Spelke, E. S., & Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, 20, 191-208.
- Balakrishnan, J. D., & Ashby, F. (1992). Subitizing: Magical numbers or mere superstition? *Psychological Research/Psychologische Forschung*, 54, 80-90.
- Bandura, A., & Walters, R. H. (1965). *Social learning and personality development*. New York, Chicago, San Francisco, Toronto, London: Holt, Rinehart and Winston.
- Banks, M. S., & Salapatek, P. (1983). Infant visual perception. In M. M. Haith & J. Campos (Eds.), *Infancy and biological development* (pp. 453-572). New York: Wiley.
- Barkow, J. H., Cosmides, L., & Tooby, J. (Eds.). (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.
- Bates, E., & Elman, J. L. (1993). Connectionism and the study of change. In M. H. Johnson (Ed.), *Brain development and cognition*. Cambridge, MA: Blackwell.
- Behr, M. J., Lesh, R., Post, T. R., & Silver, E. A. (1983). Rational-number concepts. In R. Lesh & M. Landua (Eds.), *Acquisition of mathematics concepts and processes* (pp. 91-126). New York: Academic Press.
- Berthenthal, B. I. (1993). Infants' perception of biomechanical motions: Intrinsic image and knowledge-based constraints. In C. Granrud (Ed.), *Visual perception and cognition in infancy: Carnegie Mellon symposia on cognition* (pp. 175-214). Hillsdale, NJ: Erlbaum.
- Biddell, T. R., & Fischer, K. W. (1992). Cognitive development in educational contexts: Implications for skill learning. In A. Demetriou, M. Shayer, & A. Efklides (Eds.), *Neo-Piagetian theories of cognitive development* (pp. 11-31). London and New York: Routledge & Kegan Paul.
- Bolles, R. C. (1970). Species-specific defense reactions and avoidance learning. *Psychological Review*, 77, 32-48.
- Bolles, R. C. (1988). Nativism, naturalism, and niches. In R. C. Bolles & M. D. Beecher (Eds.), *Evolution and learning* (pp. 1-15). Hillsdale, NJ: Erlbaum.
- Bornstein, M. H. (1985). Color naming versus shape naming. *Journal of Child Language*, 12, 387-393.
- Boyer, P. (1995). Causal understandings in cultural representations: Cognitive constraints on inferences from cultural input. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary approach* (pp. 615-644). Oxford, England: Oxford/Clarendon Press.

- Boysen, S. T., & Berntson, G. G. (1989). Numerical competence in a chimpanzee (*Pan troglodytes*). *Journal of Comparative Psychology*, 103, 23-31.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Brenneman, K., Massey, C. M., Machado, S., & Gelman, R. (1996). Notating knowledge about words and objects: Preschoolers' plans differ for "writing" and drawing. *Cognitive Development*, 11, 397-419.
- Broadbent, D. A. (1958). *Perception and communication*. London: Pergamon Press.
- Brown, A. L. (1990). Domain-specific principles affect learning and transfer in children. *Cognitive Science*, 14, 107-133.
- Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell & E. M. Markman (Eds.), *Cognitive development*. New York: Wiley.
- Brown, A. L., & Campione, C. (1996). Psychological theory and the design of innovative learning environments: On procedures, principles and systems. In L. Schauble & R. Glaser (Eds.), *Contributions of instructional innovation to understanding theory* (pp. 229-270). Hillsdale, NJ: Erlbaum.
- Brown, D. E. (1991). *Human universals*. Philadelphia: Temple University Press.
- Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher*, 18, 32-44.
- Bruner, J. S. (1964). The course of cognitive growth. *American Psychologist*, 19, 1-15.
- Bruner, J. S., Greenfield, P. M., Olver, et al. (1966). *Studies in cognitive development*. Cambridge, MA: Harvard University Press.
- Bullock, M., & Gelman, R. (1979). Preschool children's assumptions about cause and effect: Temporal ordering. *Child Development*, 50, 89-96.
- Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 209-253). New York: Academic Press.
- Campione, J. C., Brown, A. L., Ferrara, R., & Bryant, N. R. (1984). The zone of proximal development: Implications for individual differences and learning. *New Directions for Child Development*, 23, 77-91.
- Canfield, R. L., & Smith, E. G. (1996). Number-based expectations and sequential enumeration by 5-month-old infants. *Developmental Psychology*, 32, 269-279.
- Capaldi, E. J., & Miller, D. J. (1988). Counting in rats: Its functional significance and the independent cognitive processes which comprise it. *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 3-17.
- Carey, S. (1985). Are children fundamentally different kinds of thinkers and learners than adults? In S. Chipman, J. Segal, & R. Glaser (Eds.), *Thinking and learning skills: Research and open questions*. Hillsdale, NJ: Erlbaum.
- Carey, S. (1991). Knowledge acquisition: Enrichment or conceptual change? In S. Carey & R. Gelman (Eds.), *Epigenesis of mind: Studies in biology and cognition*. Hillsdale, NJ: Erlbaum.
- Carey, S. (1995). On the origin of causal understanding. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary approach* (pp. 268-308). Oxford, England: Oxford/Clarendon Press.
- Carey, S. (1996). Perceptual classification and expertise. In R. Gelman & T. Kit-Fong Au (Eds.), *Perceptual and cognitive development* (pp. 49-69). San Diego, CA: Academic Press.
- Carey, S., & Gelman, R. (Eds.). (1991). *The epigenesis of mind: Essays on biology and cognition*. Hillsdale, NJ: Erlbaum.
- Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 169-200). New York: Cambridge University Press.
- Case, R. (1992). Neo-Piagetian theories of intellectual development. In H. Beilin & P. B. Pufall (Eds.), *Piaget's theory: Prospects and possibilities* (pp. 21-38). Hillsdale, NJ: Erlbaum.
- Catrambone, R., & Holyoak, K. (1989). Overcoming contextual limitations on problem solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 1147-1156.
- Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition*, 23, 149-178.
- Cheng, P. W. (in press). From covariation to causation: A causal power theory. *Psychological Review*, 103.
- Chi, M. T. H. (1992). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.), *Cognitive models of science: Minnesota studies in the philosophy of science* (pp. 129-186). Minneapolis: University of Minnesota Press.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.
- Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.
- Chomsky, N. (1959). Review of the book *Verbal behavior*. *Language*, 35, 26-58.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

- Chomsky, N. (1972). Psychology and ideology. *Cognition*, 1, 11-46.
- Clark, E. V. (1993). *The lexicon in acquisition*. Cambridge, England: Cambridge University Press.
- Cole, M. (1995). Culture and cognitive development: From cross-cultural research to creating systems of cultural mediation. *Culture and Psychology*, 1, 25-54.
- Cooper, R. G., Jr. (1984). Early number development: Discovering number space with addition and subtraction. In C. Sophian (Ed.), *Origins of cognitive skill* (pp. 147-192). Hillsdale, NJ: Erlbaum.
- Cooper, R. G., Jr., Campbell, R. L., & Blevins, B. (1983). Numerical representation from infancy to middle childhood: What develops? In D. L. Rogers & J. A. Sloboda (Eds.), *The acquisition of symbolic skills* (pp. 523-533). New York: Plenum Press.
- Cosmides, L., & Tooby, J. (1994a). Origins of domain specificity: The evolution of functional organization. In L. A. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 85-116). New York: Cambridge University Press.
- Cosmides, L., & Tooby, J. (1994b). Beyond intuition and instinct blindness: Toward an evolutionarily rigorous cognitive science. *Cognition*, 50, 41-77.
- Crump, T. (1990). *The anthropology of numbers*. Cambridge, England: Cambridge University Press.
- Daly, M., & Wilson, M. (1995). Evolutionary psychology: Adaptationist, selectionist, and comparative. *Psychological Inquiry*, 6, 34-38.
- Davis, H., & Pérusse, R. (1988). Numerical competence in animals: Definitional issues, current evidence, and a new research agenda. *Behavioral and Brain Sciences*, 11, 561-615.
- Dawkins, R. (1982). *The extended phenotype: The long reach of the gene*. Oxford, England: Oxford University Press.
- Dawkins, R. (1987). *The blind watchmaker: Why the evidence of evolution reveals a universe without design*. New York: Norton.
- Dawkins, R. (1995). *River out of Eden: A Darwinian view of life*. New York: Basic Books.
- Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.
- Diamond, A. (1991). Neuropsychological insights into the meaning of object concept development. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 67-110). Hillsdale, NJ: Erlbaum.
- Dickinson, J., & Dyer, F. (1996). How insects learn about the sun's course: Alternative modeling approaches. In S. Wilson & et al. (Eds.), *From animals to animats* (Vol. 4). Cambridge, MA: MIT Press.
- Donaldson, M. (1978). *Children's minds*. New York: Norton.
- Edgerton, R. B. (1967). *The cloak of competence: Stigma in the lives of the mentally retarded*. Berkeley: University of California Press.
- Eimas, P. D. (1974). Linguistic processing of speech by young infants. In R. L. Shiefelbusch & L. L. Lloyd (Eds.), *Language perspectives: Acquisition, retardation, and intervention*. Baltimore: University Park Press.
- Elman, J. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- Flegg, G. (Ed.). (1989). *Number through the ages: Reader in the history of mathematics*. London, England: Macmillan/The Open University.
- Fodor, J. A. (1972). Some reflections on L. S. Vygotsky's *Thought and language*. *Cognition*, 1, 83-95.
- Fodor, J. A. (1975). *The language of thought*. New York: Thomas.
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
- Fodor, J. A., Garrett, M. F., Walker, E. C. T., & Parkes, C. H. (1980). Against definitions. *Cognition*, 8, 263-367.
- Fodor, J. A., & Lepore, E. (1996). The red herring and the pet fish: Why concepts still can't be prototypes. *Cognition*, 58, 253-270.
- Fuson, K. C. (1988). *Children's counting and concepts of number*. New York: Springer-Verlag.
- Gallistel, C. R. (1985). Motivation, intention and emotion: Goal-directed behavior from a cognitive-neuroethological perspective. In M. Frese & J. Sabini (Eds.), *Goal-directed behavior: The concept of action in psychology*. Hillsdale, NJ: Erlbaum.
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- Gallistel, C. R. (1992). Classical conditioning as an adaptive specialization: A computational model. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 35-67). New York: Academic Press.
- Gallistel, C. R. (1995). The replacement of general purpose theories with adaptive specializations. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1255-1267). Cambridge, MA: MIT Press.
- Gallistel, C. R., Brown, A. L., Carey, S., Gelman, R., & Keil, F. C. (1991). Lessons from animal learning for the study of cognitive development. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 3-36). Hillsdale, NJ: Erlbaum.
- Gallistel, C. R., & Gelman, R. (1990). The what and how of counting. *Cognition*, 34, 197-199.
- Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation: Numerical cognition [Special issue]. *Cognition*, 44, 43-74.
- Garcia, J., & Koelling, R. (1966). Relation of cue to consequence in avoidance learning. *Psychonomic Science*, 5, 123-124.

- Gardner, H. (1992). *Frames of mind*. New York: Basic Books.
- Gelman, R. (1972). The nature and development of early number concepts. In H. W. Reese (Ed.), *Advances in child development* (Vol. 3.). New York: Academic Press.
- Gelman, R. (1978). Cognitive development. *Annual Review of Psychology*, 29, 297-332.
- Gelman, R. (1982). Basic numerical abilities. In R. J. Stemberg (Ed.), *Advances in the psychology of intelligence* (pp. 181-205). Hillsdale, NJ: Erlbaum.
- Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive Science*, 14, 79-106.
- Gelman, R. (1991). Epigenetic foundations of knowledge structures: Initial and transcendent constructions. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 293-322). Hillsdale, NJ: Erlbaum.
- Gelman, R. (1993). A rational-constructivist account of early learning about numbers and objects. In D. Medin (Ed.), *Learning and motivation* (pp. 61-96). New York: Academic Press.
- Gelman, R. (1994). Constructivism and supporting environments. In D. Tirosh (Ed.), *Implicit and explicit knowledge: An educational approach* (pp. 55-82). Norwood, NJ: ALEX.
- Gelman, R., & Baillargeon, R. (1983). A review of some Piagetian concepts. In J. H. Flavell & E. M. Markman (Eds.), *Cognitive development* (pp. 167-230). New York: Wiley.
- Gelman, R., & Brenneman, K. (1994). First principles can support both universal and culture-specific learning about number and music. In L. A. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Culture and domain-specificity* (pp. 369-390). Cambridge, England: Cambridge University Press.
- Gelman, R., & Brown, (1986). Changing views of cognitive competence in the young. In N. Smelser & D. Gerstein (Eds.), *Discoveries and trends in behavioral and social sciences* (pp. 175-207). Commission on Behavioral and Social Sciences and Education. Washington, DC: National Research Council Press.
- Gelman, R., & Cohen, M. (1988). Qualitative differences in the way Down's Syndrome and normal children solve a novel counting problem. In L. Nadel (Ed.), *The psychobiology of Down's Syndrome* (pp. 51-99). Cambridge, MA: MIT Press.
- Gelman, R., Cohen, M., & Hartnett, P. (1989). To know mathematics is to go beyond thinking that "Fractions aren't numbers." In *Psychology of mathematics education. Volume 11 of the North American Chapter of the International Group for Psychology*. Rutgers, NJ.
- Gelman, R., Durgin, F., & Kaufman, L. (1995). Distinguishing between animates and inanimates: Not by motion alone. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary approach* (pp. 150-184). Oxford, England: Oxford/Clarendon Press.
- Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.
- Gelman, R., & Greeno, J. G. (1989). On the nature of competence: Principles for understanding in a domain. In L. B. Resnick (Ed.), *Knowing and learning: Essays in honor of Robert Glaser* (pp. 125-186). Hillsdale, NJ: Erlbaum.
- Gelman, R., & Hartnett, P. M. (in press). Early understandings of number: Paths or barriers to the construction of new understandings? *The Journal of the European Association for Learning and Instruction*, 4.
- Gelman, R., & Lee Gattis, M. (1995). Trends in educational psychology in the United States. In *Recent trends and developments in educational psychology: Chinese and American perspectives*. Paris, France: UNESCO.
- Gelman, R., & Massey, C. (1987). The cultural unconscious as contributor to the supporting environments for cognitive development. Commentary on Saxe, Guberman & Gearhart. *Society for Research in Child Development Monographs*, 52(2, Serial No. 216), 138-151.
- Gelman, R., Massey, C., & McManus, M. (1991). Characterizing supporting environments for cognitive development: Lessons from children in a museum. In J. M. Levine & L. B. Resnick (Eds.), *Socially shared cognition* (pp. 226-256). Washington, DC: American Psychological Association.
- Gelman, R., & Meck, B. (1992). Early principles aid initial but not later conceptions of number. In J. Bideaud, C. Meljac, & J. Fischer (Eds.), *Pathways to number* (pp. 171-189). Hillsdale, NJ: Erlbaum.
- Gelman, R., Spelke, E. S., & Meck, B. (1983). What preschoolers know about animate and inanimate objects: Implications for research on social cognition. In D. Rogers & J. A. Sloboda (Eds.), *The acquisition of symbolic skills* (pp. 297-326). New York: Plenum Press.
- Gelman, R., & Tucker, M. F. (1975). Further investigations of the young child's conception of number. *Child Development*, 46, 167-175.
- Gelman, S. A., & Gottfried, A. (1996). Children's causal explanations of animate and inanimate motion. *Child Development*, 67, 1970-1987.
- Gelman, S. A., & Kremer, K. E. (1991). Understanding natural cause: Children's explanations of how objects and their properties originate. *Child Development*, 62, 396-414.
- Gentner, D. (1989a). Structure mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.
- Gentner, D. (1989b). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199-241). New York: Cambridge University Press.

- Gentner, D., Rattermann, M. J., & Forbus, K. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Science*, 25, 524-575.
- Gentner, D., & Stevens, A. L. (1983). Flowing waters or teeming crowds: Mental models of electricity. In D. Gentner & A. Stevens (Eds.), *Mental models*. Hillsdale, NJ: Erlbaum.
- Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (pp. 52-77). New York: New York Academy of Sciences.
- Gibson, E. J. (1984). Perceptual development from the ecological approach. In M. C. Lamb, A. L. Brown, & B. Rogoff (Eds.), *Advances in development* (pp. 243-286). Hillsdale, NJ: Erlbaum.
- Gibson, E. J. (1988). Exploratory behavior in the development of perceiving, acting, and the acquiring of knowledge. In M. R. Rosenzweig & L. W. Porter (Eds.), *Annual review of psychology* (pp. 1-41). Palo Alto, CA: Annual Reviews.
- Gibson, E. J., Riccio, G., Stoffregen, T. A., Rosenberg, D., & Taormina, J. (1987). Detection of the transversability of surfaces by crawling and walking infants. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 533-544.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, MA: Houghton Mifflin.
- Gibson, J. J. (1983). Notes on affordances. In E. Reed & R. Jones (Eds.), *Reasons for realism* (pp. 401-418). Hillsdale, NJ: Erlbaum.
- Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review*, 62, 32-41.
- Gluck, M., & Bower, G. (1988). Evaluating an adaptive network model of human learning. *Journal of Memory and Language*, 27, 166-95.
- Gobbo, C., & Chi, M. (1986). How knowledge is structured and used by expert and novice children. *Cognitive Development*, 1, 221-237.
- Gold, E. M. (1967). Language identification in the limit. *Information and Control*, 16, 447-474.
- Golinkoff, R. M., & Harding, C. G. (1980). Infants' expectations of the movement potential of inanimate objects. In *International Conference on Infant Studies*. New Haven, CT.
- Goodman, N. (1972). *Fact, fiction, and forecast* (3rd ed.). Indianapolis, IN: Bobbs-Merrill.
- Gopnik, A., & Wellman, H. M. (1994). The theory theory. In L. A. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257-293). New York: Cambridge University Press.
- Goswami, U., & Brown, A. (1990). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition*, 35, 69-95.
- Gottlieb, G. I. E. (1983). The psychobiological approach to developmental issues. In M. M. Haith & J. J. Campos (Eds.), *Infancy and developmental psychobiology* (pp. 1-26). New York: Wiley.
- Gould, J. L., & Marler, P. (1987). Learning by instinct. *Scientific American*, 256, 74-85.
- Gould, S. J. (1991). Exaptation: A crucial tool for an evolutionary psychology. *Journal of Social Issues*, 47(3), 43-65.
- Gould, S. J., & Lewontin, R. (1979). The spandrels of San Marco and the Panglossian program: A critique of the adaptationist programme. *Proceedings of the Royal Society of London*, 250, 281-288.
- Greenfield, P. M. (1993). International roots of minority child development: Introduction to the special issue. *International Journal of Behavioral Development*, 16(3), 385-394.
- Greeno, J. G., Smith, D. R., & Moore, J. L. (1992). Transfer of situated learning. In D. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction*. Norwood, NJ: ALEX.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics: 3. Speech acts*. New York: Academic Press.
- Groen, G. J., & Resnick, L. B. (1977). Can preschool children invent addition algorithms? *Journal of Educational Psychology*, 69, 645-652.
- Haith, M. (1980). *Rules that babies look by*. Hillsdale, NJ: Erlbaum.
- Halford, G. S. (1987). A structure-mapping approach to cognitive development: The neo-piagetian theories of cognitive development: toward and integration [Special issue]. *International Journal of Psychology*, 22, 609-642.
- Halford, G. S. (1990). Is children's reasoning logical or analogical? Further comments on Piagetian cognitive developmental psychology. *Human Development*, 33(6), 356-361.
- Halford, G. S. (1993). *Children's understanding: The development of mental models*. Hillsdale, NJ: Erlbaum.
- Halford, G. S., & Wilson, W. H. (1980). A category theory approach to cognitive development. *Cognitive Psychology*, 12, 356-411.
- Hallpike, C. R. (1979). *The foundations of primitive thought*. Oxford, England: Clarendon Press.
- Hardiman, P. T., Dufresne, R., & Mestre, J. P. (1989). The relation between problem categorization and problem solving among experts and novices. *Memory and Cognition*, 17, 627-638.

- Hartnett, P. M. (1991). *The development of mathematical insight: From one, two, three to infinity*. Doctoral dissertation, University of Pennsylvania.
- Hartnett, P. M., & Gelman, R. (in press). Early understandings of number: Paths or barriers to the construction of new understandings? *The Journal of the European Association for Learning and Instruction*, 4.
- Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology*, 108, 356–388.
- Hasher, L., & Zacks, R. T. (1984). Automatic processing of fundamental information: The case of frequency of occurrence. *American Psychologist*, 39, 1372–1388.
- Held, R. (1993). What can rates of development tell us about underlying mechanisms? In C. Granrud (Ed.), *Visual perception and cognition in infancy: Carnegie Mellon symposia on cognition* (pp. 75–89). Hillsdale, NJ: Erlbaum.
- Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature*, 370, 57–59.
- Hinde, R. A., & Stevenson-Hinde, J. S. (Eds.). (1973). *Constraints on learning*. New York: Academic Press.
- Hirschfeld, L. A. (1996). *Race in the making*. Cambridge, MA: MIT Press.
- Hirschfeld, L. A., & Gelman, S. A. (Eds.). (1994). *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge University Press.
- Holyoak, K., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.
- Hull, C. L. (1952). *A behavior system*. New Haven, CT: Yale University Press.
- Hurtford, J. R. (1987). *Language and number*. Cambridge, England: Cambridge University Press.
- Ibrahim, G. (1985). *From one to zero*. New York: Viking.
- Inagaki, K., & Hatano, G. (1987). Young children's spontaneous personification as analogy. *Child Development*, 58, 1013–1020.
- Inhelder, B., & Piaget, J. (1956). *The growth of logical thinking from childhood to adolescence*. New York: Basic Books.
- Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child*. New York: Harper & Row.
- Jeyitous, S. W. (1986). *Altmodemo: Semantic conceptual development among the Yoruba*. Doctoral dissertation, Cornell University, Ithaca, NY.
- Johnson, M. H., & Gilmore, R. O. (1996). Developmental cognitive science: A biological perspective on cognitive change. In R. Gelman & T. Kit-Fong Au (Eds.), *Perceptual and cognitive development: Handbook of perception and cognition* (pp. 333–372). San Diego, CA: Academic Press.
- Johnson, M. H., & Morton, J. (1991). *Biology and cognitive development: The case of face recognition*. Cambridge, England: Blackwell.
- Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subadjacency in the acquisition of a second language. *Cognition*, 39, 215–258.
- Kail, R., & Bisanz, J. (1992). The information-processing perspective on cognitive development. In R. J. Sternberg & C. A. Berg (Eds.), *Intellectual development*. New York: Cambridge University Press.
- Kalnins, I. V., & Bruner, J. S. (1973). The coordination of visual observation and instrumental behavior in early infancy. *Perception*, 2, 307–314.
- Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.
- Karmiloff-Smith, A., & Inhelder, B. (1974–1975). If you want to get ahead, get a theory. *Cognition*, 3, 195–212.
- Kehoe, E. J. (1988). A layered network model of associative learning: Learning to learn and configuration. *Psychological Review*, 95, 411–433.
- Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, 88, 197–227.
- Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT.
- Keil, F. C. (1990). Constraints on constraints. Surveying the epigenetic landscape. *Cognitive Science*, 14, 135–168.
- Keil, F. C. (1994). The birth and nurturance of concepts by domains: The origins of concepts of living things. In L. A. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 234–254). Cambridge, England: Cambridge University Press.
- Keil, F. C. (1995). The growth of causal understandings of natural kinds. In D. Sperber, D. P. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary approach* (pp. 234–267). Oxford, England: Oxford/Clarendon Press.
- Keliman, P. (1996). The origins of object perception. In R. Gelman & T. Kit-Fong Au (Eds.), *Perceptual and cognitive development* (pp. 3–48). New York: Academic Press.
- Kemler, D. (1983). Holistic and analytic modes in perceptual and cognitive development. In T. Tighe & B. E. Shepp (Eds.), *Perception, cognition, and development: Interactional analyses*. Hillsdale, NJ: Erlbaum.
- Kitcher, P. (1982). *The nature of mathematical knowledge*. New York: Oxford University Press.
- Klahr, D., & Wallace, J. G. (1973). The role of quantification operators in the development of conservation of quantity. *Cognitive Psychology*, 4, 301–327.

- Koslowski, B. (1996). *The development of scientific reasoning*. Cambridge, MA: MIT Press.
- Koslowski, B., Spilton, D., & Snipper, A. (1981). Children's beliefs about instances of mechanical and electrical causation. *Journal of Applied Developmental Psychology*, 2, 189-210.
- Koslowski, B., & Winsor, A. P. (1981). *Preschool children's spontaneous explanations and requests for explanations: A non-human application of the child-as-scientist metaphor*. Unpublished manuscript, Department of Human Development, Cornell University, Ithaca, NY.
- Kuhn, D. (1996). *Microgenetic study of change. What has it told us?* Unpublished manuscript, Psychology, Teacher's College, Columbia University, New York.
- Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development*, 60(4, Serial No. 241), 1-158.
- Kuhn, T. S. (1970). *The structure of scientific resolutions* (2nd ed.). Chicago: University of Chicago Press.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91-196). Cambridge, England: Cambridge University Press.
- Lampert, M. (1986). Knowing, doing, and teaching multiplication. *Cognition and Instruction*, 3, 305-342.
- Landau, B., & Gleitman, L. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.
- Larkin, J. H., McDermott, J., Simon, D., & Simon, H. A. (1980). Models of competence in solving physics problems. *Cognitive Science*, 4, 317-345.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, MA: Cambridge University Press.
- Lee, K., & Karmiloff-Smith, A. (1996). The development of external symbol systems: The child as a notator. In R. Gelman & T. Kit-Fong Au (Eds.), *Perceptual and cognitive development* (pp. 185-211). San Diego, CA: Academic Press.
- Lehrman, D. S. (1970). Semantic and conceptual issues in the nature-nurture problem. In L. Aaronson, E. Tobach, D. S. Lehrman, & J. S. Rosenblatt (Eds.), *Development and evolution of behavior: Essays in memory of T. C. Schneirla* (pp. 17-24). San Francisco: Freeman.
- Leinhardt, G., Putnam, R., & Harttrup, R. A. (Eds.). (1992). *Analysis of arithmetic for mathematics teaching*. Hillsdale, NJ: Erlbaum.
- Leslie, A. M. (1988). Some implications of pretense for mechanisms underlying the child's theory of mind. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 19-46). Cambridge, England: Cambridge University Press.
- Leslie, A. M. (1994a). Pretending and believing: Issues in the theory of ToMM. *Cognition*, 50, 211-238.
- Leslie, A. M. (1994b). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119-148). New York: Cambridge University Press.
- Leslie, A. M. (1995). A theory of agency. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary approach* (pp. 121-141). Oxford, England: Oxford/Clarendon Press.
- Leslie, M., & Keebie, S. (1987). Do six-month-old infants perceive causality? *Cognition*, 25, 265-288.
- Locke, J. (1690). *An essay concerning human understanding*.
- Loftus, E. F. (1991). Made in memory: Distortions in recollection after misleading information. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 27, pp. 187-215). San Diego, CA: Academic Press.
- Lorenz, K. (1966). *Evolution and modification of behavior*. London: Methuen.
- Macario, J. F. (1991). Young children's use of color and classification: Foods and canonically colored objects. *Cognitive Development*, 6, 17-46.
- Mandler, J. M. (1996). Development of categorization: Perceptual and conceptual categories. In G. Bremner, A. Slater, & G. Butterworth (Eds.), *Infant development: Recent advances*. Hove, England: Erlbaum.
- Mandler, J. M., & McDonough, L. (1993). Concept formation in infancy. *Cognitive Development*, 8, 291-318.
- Mandler, J. M., & McDonough, L. (1996). Drinking and driving don't mix: Inductive generalization in infancy. *Cognition*, 59, 307-335.
- Mandler, J. M., & Orlich, F. (1993). Analogical transfer: The roles of schema abstraction and awareness. *Bulletin of the Psychonomic Society*, 31, 485-487.
- Marcus, G. F. (1996). Why do children say "broke"? *Current Directions in Psychological Science*, 5(3), 81-85.
- Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., & Xu, F. (1992). Over regularization in language acquisition. *Monographs of the Society for Research in Child Development*, 57(4, Serial No. 228), 1-182.
- Markman, E. M. (1989). *Categorization and naming in children*. Cambridge, MA: MIT Press.
- Marler, P. (1991). The instinct to learn. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind* (pp. 37-66). Hillsdale, NJ: Erlbaum.

- Massey, C. M. (1989). The development of the animate-inanimate distinction in preschoolers. *Dissertation Abstracts International*, 50(n2-B), 766.
- Massey, C. M., & Gelman, R. (1988). Preschoolers decide whether pictured unfamiliar objects can move themselves. *Developmental Psychology*, 24, 307-317.
- Mayr, E. (1983). How to carry out the adaptationist program. *The American Naturalist*, 121, 324-334.
- McCleish, J. (1992). *The history of numbers and how they shape our lives*. New York: Fawcette Columbine.
- McClelland, J., & Rumelhart, D. (1986). *Parallel distributed processing* (Vol. 2). Cambridge, MA: MIT Press.
- McCloskey, M., & Cohen, N. J. (1989). Catastrophic interference in connectionist networks: The sequential learning problem. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 109-165). New York: Academic Press.
- McShane, J. (1991). *Cognitive development: An information process approach*. Cambridge, MA: Basil Blackwell.
- Meck, W. H., & Church, R. M. (1984). The numerical attribute of stimuli. In H. L. Roitblatt, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 445-464). Hillsdale, NJ: Erlbaum.
- Medin, D. L. (1989). Concepts and conceptual structure. *American Psychologist*, 44, 1469-1489.
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review*, 100, 254-278.
- Mehler, J., & Christophe, A. (1995). Maturation and learning of language in the first year of life. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 943-954). Cambridge, MA: MIT Press.
- Miller, K. F. (1996). Origins of quantitative competence. In R. Gelman & T. Kit-Fong Au (Eds.), *Perceptual and cognitive development* (pp. 213-241). San Diego, CA: Academic Press.
- Miller, K. F., Smith, C. M., Zhu, J., & Zhang, H. (1995). Preschool origins of cross-national differences in mathematical competence: The role of number-naming systems. *Psychological Science*, 6, 56-60.
- Miller, N. E., & Dollard, J. (1941). *Social learning and imitation*. New Haven, CT: Yale University Press.
- Morris, R. G. M. (1981). Spatial localization does not require the presence of local cues. *Learning and Motivation*, 12, 239-260.
- Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-month-old infants. *Cognition*, 47, 121-148.
- Neisser, U. (1966). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive Psychology*, 7, 480-494.
- Nelson, D. A., & Marler, P. (1993). Innate recognition of song in white-crowned sparrows: A role in selective vocal learning? *Animal Behaviour*, 46, 806-808.
- Nelson, K. (1978). How children represent knowledge of their world in and out of language: A preliminary report. In R. S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, NJ: Erlbaum.
- Nelson, K. (1987). Nativist and functionalist views of cognitive development: Reflections on Keil's review of "Making sense: The acquisition of shared meaning." *Cognitive Development*, 2, 237-247.
- Nelson, K. (1988). Constraints on word learning? *Cognitive Development*, 3, 221-246.
- Nesher, P., & Sukenik, M. (1991). The effects of formal representation on the learning of rational concepts. *Learning and Instruction*, 1, 161-175.
- Newell, A., Simon, H. A., & Shaw, (1958). Elements of a theory of human problem solving. *Psychological Review*, 65, 151-166.
- Newport, E. L. (1990). Maturation constraints on language learning. *Cognitive Science*, 14, 11-28.
- Newport, E. L., & Supalla, T. (1990). *A critical period effect in the acquisition of a primary language*. Unpublished manuscript, University of Rochester, Rochester, NY.
- Nunes, T., & Bryant, P. (1996). *Children doing mathematics*. Cambridge, MA: Blackwell.
- Nunes, T., Schliemann, A. D., & Carraher, D. W. (1993). *Street mathematics and school mathematics*. Cambridge, England: Cambridge University Press.
- Ornstein, P. A., & Naus, M. J. (1978). Rehearsal processes in children's memory. In P. A. Ornstein (Ed.), *Memory development in children*. Hillsdale, NJ: Erlbaum.
- Osherson, D. N., Stob, M., & Weinstein, S. (1986). *Systems that learn: An introduction to learning theory for cognitive and computer scientists*. Cambridge, MA: MIT Press.
- Parisi, D. (1996). Computational models of developmental mechanisms. In R. Gelman & T. Kit-Fong Au (Eds.), *Perceptual and cognitive development* (pp. 373-412). San Diego, CA: Academic Press.
- Pearce, J. M. (1994). Discrimination and categorization. In N. J. Mackintosh (Ed.), *Animal learning and cognition: Handbook of perception and cognition* (2nd ed., pp. 109-134). San Diego, CA: Academic Press.
- Pepperberg, I. M. (1987). Evidence for conceptual quantitative abilities in the African grey parrot: Labeling of cardinal sets. *Ethology*, 75, 37-61.
- Piaget, J. (1929). *The child's conception of the world*. London, England: Routledge & Kegan Paul.
- Piaget, J. (1952). *The child's conception of number*. London, England: Routledge & Kegan Paul.

- Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology*. New York: Wiley.
- Piaget, J. (1971). *Biology and knowledge*. Chicago: University of Chicago Press.
- Piaget, J., & Inhelder, B. (1956). *The growth of logical thought*. London, England: Routledge & Kegan Paul.
- Piattelli-Palmarini, M. (1989). Evolution, selection, and cognition: From "learning" to parameter setting in biology and in the study of language. *Cognition*, 31, 1-44.
- Pinker, S. (1991). Rules of language. *Science*, 253, 530-544.
- Platt, J. R., & Johnson, D. M. (1971). Localization of position within a homogeneous behavior chain: Effects of error contingencies. *Learning and Motivation*, 2, 386-414.
- Poulin-Dubois, D., & Shultz, T. (1989). The infant's concept of agency: The distinction between social and nonsocial objects. *Journal of Genetic Psychology*, 151(1), 77-90.
- Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, 36, 1-16.
- Premack, D. (1996). Piggyback with Darwin. *Contemporary Psychology*, 41, 207-212.
- Premack, D., & Premack, A. J. (1995). Intention as psychological cause. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 185-199). New York: Clarendon Press/Oxford University Press.
- Quine, Q. V. (1960). *Word and object*. Cambridge, MA: MIT Press.
- Rabinowicz, T. (1979). The differentiate maturation of the human cerebral cortex. In F. Falkner & J. M. Tanner (Eds.), *Human growth: Vol. 3. Neurobiology and nutrition*. New York: Plenum Press.
- Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, 331, 163-166.
- Reed, H. J., & Lave, J. (1979). Arithmetic as a tool for investigating relations between culture and cognition. *American Ethnologist*, 6, 568-582.
- Resnick, L. B. (1989). Developing mathematical knowledge. *American Psychologist*, 44, 162-169.
- Resnick, L. B. (1994). Performance puzzles: Educational reform through national standards and assessment [Special issue]. *American Journal of Education*, 102, 511-526.
- Rilling, M. (1967). Number of responses as a stimulus in fixed interval and fixed ratio schedules. *Journal of Comparative and Physiological Psychology*, 63, 60-65.
- Rogoff, B., & Chavayah, P. (1995). What's become of research on the cultural basis of cognitive development? *American Psychologist*, 50, 859-877.
- Rogoff, B., & Lave, J. (1984). *Everyday cognition: Its development in social context*. Cambridge, MA: Harvard University Press.
- Rosch, E. H., Mervis, C. B., Gray, W. D., Boyes-Braem, P., & Johnson, D. N. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.
- Rosser, R. (1991). *Cognitive development: Psychological and biological perspectives*. Boston: Allyn & Bacon.
- Rovee-Collier, C., & Bhatt, R. S. (1993). Evidence of long-term memory in infancy. In R. Vasta (Ed.), *Annals of child development* (pp. 1-45). London, England: Jessica Kingsley.
- Rozin, P. (1976). The evolution of intelligence and access to the cognitive unconscious. In J. A. Sprague & A. N. Epstein (Eds.), *Progress in psychobiology and physiological psychology* (pp. 245-280). New York: Academic Press.
- Rozin, P., & Kalat, J. W. (1971). Specific hungers and poison avoidance as adaptive specializations of learning. *Psychological Review*, 78, 459-486.
- Rozin, P., & Schull, J. (1988). The adaptive-evolutionary point of view in experimental psychology. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Learning and cognition* (pp. 503-546). New York: Wiley.
- Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986). *Parallel distributed processing*. Cambridge, MA: MIT Press.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.
- Saxe, G. B. (1979). Developmental relations between notational counting and number conservation. *Child Development*, 50, 180-187.
- Saxe, G. B. (1981). Body parts as numerals: A developmental analysis of numeration among the Oksapmin in Papua, New Guinea. *Child Development*, 52, 306-316.
- Saxe, G. B. (1990). *Culture and cognitive development: Studies in mathematical understanding*. Hillsdale, NJ: Erlbaum.
- Scarr, S. (1993). Genes, experience, and development. In D. Magnusson, P. Jules, & M. Casaer (Eds.), *Longitudinal research on individual development: Present status and future perspectives. European network on longitudinal studies on individual development* (pp. 26-50). Cambridge, England: Cambridge University Press.
- Schieffelin, B. B., & Ochs, E. (1986). Language socialization. *Annual Review of Anthropology*, 15.
- Schwartz, B., & Reisberg, D. (1991). *Learning and memory*. New York: Norton.
- Shatz, M. (1978). The relationship between cognitive processes and the development of communication skills. In C. B. Keasey (Ed.), *Nebraska Symposium on Motivation* (pp. 1-42). Lincoln: University of Nebraska Press.
- Shatz, M., & Backscheider, A. (1991, November). *Acquiring the normative concepts of color and number*. Presented at the meeting of the Psychonomic Society, Seattle, WA.
- Shepard, R. N. (1987). Evolution of a mesh between principles of the mind and regularities of the world. In J. Dupré (Ed.), *The*

- latest on the best: *Essays on evolution and optimality*. Cambridge, MA: MIT Press.
- Shepard, R. N. (1994). Perceptual-cognitive universals as reflections of the world. *Psychonomic Bulletin and Review*, 1, 2-28.
- Shettleworth, S. J. (1993). Varieties of learning and memory in animals. *Journal of Experimental Psychology: Animal Behavior Processes*, 19, 5-14.
- Siegal, M. (1991). *Knowing children: Experiments in conversations and cognition*. Hove and London, England: Erlbaum.
- Siegal, M. (1995). Becoming mindful of food and conservation. *Current Directions*, 4, 177-181.
- Siegal, M. (1996). Conversation and cognition. In R. Gelman & T. Kit-Fong Au (Eds.), *Perceptual and cognitive development* (pp. 185-211). San Diego, CA: Academic Press.
- Siegler, R. S. (1983). Information processing approaches to cognitive development. In W. Kessen (Ed.), *History, theory and methods*. New York: Wiley.
- Siegler, R. S. (1991). *Children's thinking* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Science*, 28, 225-273.
- Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist*, 46, 606-620.
- Siegler, R. S., & Crowley, K. (1994). Constraints on learning in nonprivileged domains. *Cognitive Psychology*, 27, 194-226.
- Siegler, R. S., & Robinson, M. (1982). The development of numerical understanding. In H. W. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior*. New York: Academic Press.
- Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon & G. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). Hillsdale, NJ: Erlbaum.
- Silver, E. A., Mamona-Downs, J., Leung, S. S., & Kenny, P. A. (1996). Posing mathematical problems: An exploratory study. *Journal for Research in Mathematics Education*, 27, 293-309.
- Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. S. Siegler (Ed.), *Children's thinking: What develops?* Hillsdale, NJ: Erlbaum.
- Simon, H. A. (1972). On the development of the processor. In S. Farnham-Diggory (Ed.), *Information processing in children*. New York: Academic Press.
- Simon, H. A. (1981). *The sciences of the artificial*. Cambridge, MA: MIT Press.
- Simon, T. J., & Klahr, D. (1995). A computational theory of children's learning about conservation. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 315-354). Hillsdale, NJ: Erlbaum.
- Simons, D. J., & Keil, F. C. (1995). An abstract to concrete shift in the development of biological thought: The inside story. *Cognition*, 56, 129-163.
- Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review*, 57, 193-216.
- Slotta, J. D., & Chi, M. T. H. (1996). *Understanding constraint-based processes: A precursor to conceptual change in physics*. Cognitive Science Society. La Jolla, CA: Erlbaum.
- Slotta, J. D., Chi, M. T. H., & Joram, E. (1995). Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change. *Cognition and Instruction*, 1(3), 373-400.
- Sophian, C. (1994). *Children's numbers*. Madison, WI: WCB Brown & Benchmark.
- Sophian, C., & Adams, N. (1987). Infants' understanding of numerical transformations. *British Journal of Developmental Psychology*, 5, 257-264.
- Spelke, E. S. (1990). Principles of object perception. *Cognitive Science*, 14, 29-56.
- Spelke, E. S. (1991). Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition. The Jean Piaget Symposium series* (pp. 133-169). Hillsdale, NJ: Erlbaum.
- Spelke, E. S., & Hermer, L. (1995). Early cognitive development: Objects and space. In R. Gelman & T. Kit-Fong Au (Eds.), *Perceptual and cognitive development* (pp. 72-114). San Diego, CA: Academic Press.
- Spelke, E. S., Phillips, A., & Woodward, A. L. (1995). Infants' knowledge of object motion and human action. In D. Sperber, D. P. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 44-78). Oxford, England: Clarendon Press/Oxford University Press.
- 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., Premack, D., & A. J. Premack (Eds.), (1995). *Causal cognition: A multidisciplinary approach*. Oxford, England: Clarendon Press/Oxford University Press.
- Starkey, P., & Cooper, R. G. (1995). The development of subitizing in young children. *British Journal of Developmental Psychology*, 13, 399-420.
- Starkey, P., & Gelman, R. (1982). The development of addition and subtraction abilities prior to formal schooling. In T. P.

- Carpenter, J. M., Moser, & T. A. Romberg (Eds.). *Addition and subtraction: A developmental perspective*. Hillsdale, NJ: Erlbaum.
- Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition*, 36, 97-127.
- Stern, D. (1977). *The first relationship: Infant and mother*. Cambridge, MA: Harvard University Press.
- Stevenson, H. W., Lee, S. L., & Stigler, J. W. (1986). Mathematics achievement of Chinese, Japanese, and American children. *Science*, 231, 693-699.
- Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap*. New York: Summit Books.
- Stigler, J. W., & Fernandez, C. (1995). Learning mathematics from classroom instruction: Cross-cultural and experimental perspectives. In C. A. Nelson (Ed.), *Basic and applied perspectives on learning, cognition, and development. The Minnesota Symposia on Child Psychology* (Vol. 28, pp. 103-130). Mahwah, NJ: Erlbaum.
- Symons, D. (1990). Adaptiveness and adaptation. *Ethology & Sociobiology*, 11, 427-444.
- Thorndike, E. L. (1913). *Educational psychology*. New York: Teacher's College, Columbia University.
- Thurlow, M. L., & Turner, J. E. (1977). Children's knowledge of money and time: Effective instruction for the mentally retarded. *Education and Training of the Mentally Retarded*, 203-212.
- Tolchinsky-Landsmann, L. (1990). Early writing development: Evidence from different orthographic systems. In M. Spoolers (Ed.), *Literacy acquisition*. Norwood, NJ: ALEX.
- Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge, England: Cambridge University Press.
- Tooby, J., & Cosmides, L. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture*. New York: Oxford University Press.
- Trabasso, T., & Bower, G. H. (with Gelman, R.). (1968). *Attention in learning*. New York: Wiley.
- Treisman, A. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 194-214.
- Tronick, E. Z. (1989). Emotions and emotional communication in infants. *American Psychologist*, 44, 112-119.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.
- VanderWall, S. B. (1982). An experimental analysis of seed recovery in Clark's nutcracker. *Animal Behavior*, 30, 84-94.
- Vergnaud, G. (1983). Multiplicative structures. In R. Lesh & M. Landau (Eds.), *Acquisition of mathematics concepts and processes*. New York: Academic Press.
- Vosniadou, S., & Ortony, A. (1989). *Similarity and analogical reasoning*. New York: Cambridge University Press.
- Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press.
- Watson, J. B. (1913). Psychology as the behaviorist views it. *Psychological Review*, 20, 158-177.
- Watson, J. S. (1987). Reactions to response-contingent stimulation in early infancy. In J. Oates & S. Sheldon (Eds.), *Cognitive development in infancy*. Hove, England: Erlbaum.
- Waxman, S. (1990). Convergences between lexical and conceptual organization. In J. P. Byrnes & S. A. Gelman (Eds.), *Perspectives on language and cognition: Interrelations in development*. Cambridge, England: Cambridge University Press.
- Wellman, H. M., & Gelman, S. A. (1992). Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, 43, 337-375.
- White, B. Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction*, 10, 1-100.
- Williams, E. M., & Gelman, R. (1995, April). *Preschoolers' theory of action: Causal accounts of why familiar animate and inanimate objects move and stop*. Poster presented at the biennial meeting of the Society for Research in Child Development, Indianapolis, IN.
- Wilson, D., & Sperber, D. (1986). *Relevance*. Oxford, England: Blackwell.
- Wiser, M. (1987). Novice and historical thermal theories. In S. Strauss (Ed.), *Ontogeny, phylogeny, and the history of science*. Norwood, NJ: ALEX.
- Wittgenstein, L. (1953). *Philosophical investigations*. New York: Macmillan.
- Wynn, K. (1990). Children's understanding of counting. *Cognition*, 36, 155-193.
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358, 749-750.
- Wynn, K. (1995). Infants possess a system of numerical knowledge. *Current Directions in Psychological Science*, 4, 172-177.
- Wynn, K. (1996). Infants' individuation and enumeration of actions. *Psychological Science*, 7, 164-169.
- Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology*, 30, 111-153.
- Young, J. Z. (1957). *The life of mammals*. Oxford, England: Oxford University Press.
- Zaslavsky, C. (1973). *Africa counts*. Boston: Prindle, Wever & Schmidt.

