Piaget also described children in terms of general intellectual traits, such as egocentrism. These trait descriptions fit young children's thinking in many ways, but not in all. For example, although 5-year-olds are egocentric in some situations, they and even younger children behave nonegocentrically in other situations. Moreover, even older children and adults sometimes behave egocentrically. The trait descriptions thus seem to be in the right ballpark, but to gloss over exceptions. More generally, Piaget's theory continues to be of interest because it communicates a good feel for children's thinking and because it asks the right questions.

Recommended Readings

Brainerd, C.J. (Ed.) (1996). PS celebrates the centennial of Jean Piaget [special section]. *Psychological Science, 7*(4). This special issue was organized to honor Piaget on the 100th anniversary of his birth. The articles provide an overview of his life, his contributions and his legacies to the field of cognitive development.


Scene: Daughter and father in their yard. A playmate rides in on a bike.

CHILD: Daddy, would you unlock the basement door?

FATHER: Why?

C: Cause I want to ride my bike.

F: Your bike is in the garage.

C: But my socks are in the dryer. (Klahr, 1978, pp. 181–182)

What thinking underlay this child's enigmatic comments? David Klahr, a prominent information-processing theorist, built the following model of the thinking that led to her initial request for him to unlock the basement door:

Top goal: I want to ride my bike.

Constraint: I need shoes to ride comfortably.

Fact: I'm barefoot.

Subgoal 1: Get my sneakers.

Fact: The sneakers are in the yard.

Fact: They're uncomfortable on bare feet.

Subgoal 2: Get my socks.

Fact: The sock drawer was empty this morning.

Inference: The socks probably are in the dryer.
As this example suggests, the information-processing approach to development speaks to the essential tension within children’s thinking, the tension produced by children ceaselessly striving to reach their goals despite patchy knowledge, limited processing capacity, and obstacles posed by the external world. The particular strategy used in the story was means-ends analysis, which involves repeatedly comparing one’s current state with one’s goal and then taking steps to reduce the distance between them. In other situations, children use other strategies. To overcome their limited memory capacities, they use strategies such as rehearsal (repeating material over and over before recalling it, as when trying to remember a phone number). To overcome their limited knowledge, they use the tools provided by the culture in which they live: dictionaries, encyclopedias, calculators, the Internet, older children and adults who will answer their questions, and other devices and resources.

Information-processing theories of development vary among themselves, but they all share several basic assumptions. The most fundamental assumption is that thinking is information processing. Rather than focusing on stages of development, they focus on the information that children represent, the processes that they apply to the information, and the memory limits that constrain the amount of information they can represent and process. Cognitive growth is analyzed in terms of age-related and experience-related changes in these capabilities. Information-processing analyses generally are more precise than those of stage approaches; the detailed analysis of goals, subgoals, knowledge, and inferences within Klahr’s model of his daughter’s thinking is characteristic.

A second defining characteristic of information-processing theories of development is an emphasis on precise analysis of change mechanisms. Two critical goals are to identify the change mechanisms that contribute most to development and to specify exactly how these change mechanisms work together to produce cognitive growth. The flip side of this emphasis on how development occurs is an emphasis on the cognitive limits that prevent development from occurring more rapidly than it does. Thus, information-processing theories attempt to explain both how children of given ages have come as far as they have and why they have not gone further.

A third assumption of most information-processing approaches is that change is produced by a process of continuous self-modification. That is, the outcomes generated by the child’s own activities change the way the child will think in the future. For example, in Shrage and Siegler’s (1998) model of strategy choice, use of alternative strategies creates increasing knowledge concerning the effectiveness of each strategy which, in turn, changes the strategies that are used. Such self-modifying processes eliminate the need to account for special age-defined transition periods, as in Piaget’s proposed transition from the concrete operations to the formal operations stage around age 12. Instead, children’s thinking is viewed as continuously changing at all ages.

What is the relation of information-processing approaches to alternative views, such as the Piagetian approach? The two approaches have quite a bit in common. Both are aimed at answering the same fundamental questions: “What develops?” and “How does development occur?” Both try to identify children’s cognitive capabilities and limits at various points in development. Both try to explain how later, more advanced understandings grow out of earlier, more primitive ones.

The two approaches also differ in important ways, though. Information-processing approaches place greater emphasis on the role of processing limitations, strategies for overcoming the limitations, and knowledge about specific content. There also is a greater emphasis on precise analyses of change and on the contribution of ongoing cognitive activity to that change. These differences have led to a greater use of formalisms, such as computer simulations and flow diagrams, which allow information-processing theorists to model in detail how thinking proceeds. Because of this focus on precise specification of cognitive processes, information-processing accounts of cognitive development often involve detailed, in-depth analyses of children’s performance on a single task or on a narrow range of tasks. In contrast, the Piagetian approach seeks to characterize children’s thinking across a broad range of tasks and content domains.

A final difference is that information-processing theories assume that our understanding of how children think can be greatly enriched by knowledge of how adults think. The underlying belief is that just as we can more deeply understand our own adult thinking when we appreciate how it developed, we can also better understand the development of children’s thinking when we know where the development is going.

This chapter is divided into two main sections. In the first, we examine the basic information-processing framework. This framework provides a way of thinking about the cognitive systems of both children and adults. In the second main section, we consider five information-processing theories that focus on development. No one of these theories covers the huge expanse of topics and ages encompassed by Piaget’s theory. On the other hand, each provides more precise and complete characterizations of particular aspects of development than Piaget did. The chapter’s organization is outlined in Table 3.1.
TABLE 3.1 Chapter Outline

I. An Overview of the Information-Processing System
   A. Structural Characteristics
   B. Processes
II. Information-Processing Theories of Development
   A. Neo-Piagetian Theories
   B. Psychometric Theories
   C. Production System Theories
   D. Connectionist Theories
   E. Theories of Cognitive Evolution
III. Summary

An Overview of the Information-Processing System

Any cognitive theory must come to grips with two basic characteristics of human cognition. First, our thinking is limited, both in the amount of information that we can attend to simultaneously and in the speed with which we can process the information. Second, our thinking is flexible, capable of adapting to constantly changing goals, circumstances, and task demands. Information-processing theories have attempted to come to grips with this dual nature of cognition by focusing on both structural characteristics, which determine the limits within which thinking occurs, and processes, which provide the means for flexible adaptation to a constantly changing world.

 Structural Characteristics

Structural characteristics of the information-processing system provide its basic organization. They sometimes are referred to as the cognitive architecture; the analogy is to the architectural plan for a building, which specifies its main characteristics, but not the more detailed features. Structural features of the cognitive system tend to be relatively enduring; the same basic organization is believed to be maintained throughout development. Structural features are also universal; all children have the same basic cognitive organization, though the efficiency with which the different parts operate varies across individuals and age groups. This basic organization is often viewed within a three-part framework: sensory memory, working memory, and long-term memory.

Sensory memory. People possess a special capacity for briefly retaining relatively large amounts of information that they have just encountered. This capacity is often labeled sensory memory. Sperling (1960) established several characteristics of sensory memory that influence processing of visual information. He presented college students a three-by-four matrix of letters for one-twentieth of a second. When asked immediately after the presentation to name the letters, the college students typically recalled four or five, about 40 percent of the list. Then Sperling changed the procedure in a small but important way. Rather than having the students recall all of the letters, he asked them to recall only the letters in one row. Since it was impossible to anticipate the identity of the row, the students needed to process all 12 letters, just as in the original task. However, requiring them to recite the contents of only one row eliminated their need to retain the information while they named the first few letters.

Sperling found that when the experimenter indicated which row to recall immediately after the display was shut off, the college students recalled 80 percent of the letters in the row. When the row's identity was indicated one-third of a second after the display was turned off, their recall declined to 55 percent. When it was indicated one second after, performance declined to the original 40 percent. Sperling's interpretation was that a one-twentieth-second exposure was sufficient to enable participants to create a visual icon (a literal copy of the original stimulus), but that the icon faded within one-third of a second and disappeared after a second. These estimates remain reasonable in light of subsequent research.

The capacity of children's sensory memory appears to increase with development. Cowan and colleagues (Cowan, Nugent, Elliott, Ponomarev, & Saults, 1999) investigated this issue in a study of sensory memory for auditory information. Participants were asked to play an attention-demanding computer game, while at the same time listening to lists of spoken digits with short intervals of silence in between the lists. Occasionally (roughly once every 13 lists), participants received a cue to report the most recent list of digits. Cowan and colleagues reasoned that, under these conditions, participants would need to recall the list from their auditory sensory memory.

Cowan and colleagues tested first-grade students, fourth-grade students, and adults using this paradigm. On average, first graders recalled about 2.5 digits, fourth graders recalled about 3 digits, and adults recalled about 3.5 digits. Thus, the capacity of children's sensory memory appears to increase with age.

Working memory. Working memory is where active thinking occurs: constructing new strategies, computing solutions to arithmetic problems, comprehending what we read, and so on. Its operation involves combining information coming into sensory memory with information stored in long-term memory and transforming that information into new forms. For example, when we read a book, working memory combines the sensory information about the words on the page with long-term memory representations of the meanings of the words, and uses both sources of data to represent the meaning of the text as a whole.

The operation of working memory is limited in several ways. The first is its capacity, which is the number of units it can operate on at one time. This number is not large; it is usually estimated to be between three and seven units. Being more precise than this is difficult, because the exact estimates depend on the particulars of the task on which the capacity is measured. For example,
estimates of capacity tend to be larger when they are based on the number of numbers that can be maintained in memory than when based on the number of letters that can be maintained (Dempster, 1981).

The limit on working-memory capacity is a limit on the number of meaningful units (chunks) that can be operated on, rather than on the number of physical units. A letter, a number, a word, or a familiar phrase can function as a single chunk, because each is a single unit of meaning. Thus, it is as easy to remember a set of three unrelated words with nine letters (hit, red, cup) as to remember three unrelated letters (g, f, r) (Miller, 1956).

The rate at which information is lost from working memory also limits cognitive functioning. Material ordinarily is lost within 15 to 30 seconds. However, at least with verbally encoded information such as words or numbers, rehearsal can maintain the information in working memory for a longer time.

Older children can maintain considerably more information in working memory than can younger ones. A large part of the reason appears to be the older children's more rapid rate of rehearsal. In general, the faster that both adults and children can rehearse verbal material, the more material they can maintain in working memory (Baddeley, 1986; Baddeley & Hitch, 1974). Faster rehearsal means less time between repetitions of a given word, and thus less likelihood that the word will be forgotten before it is rehearsed again. As shown in Figure 3.1, rate of pronunciation of words is closely related to the number of words that can be maintained in working memory. Older children's greater speed of pronunciation appears to be a large part of the reason why they can maintain more material in working memory (Hitch & Towse, 1995).

Working memory appears to include separate storage capacities for verbal and spatial information, together with an executive processor that controls attention to different sources of information (Baddeley, 1986). Information from the verbal and spatial subsystems is integrated and coordinated with information from long-term memory in a workspace termed the "episodic buffer" (Baddeley, 2000).

The development of working memory appears to involve both changes in the amount of verbal and spatial information of each type that can be remembered, and increasingly effective separation between the two. Evidence comes from a study in which 8-year-olds, 10-year-olds and college students were presented either a series of digits, which would usually be coded verbally, or a series of locations of Xs on a tic-tac-toe grid, which usually would be coded spatially (Hale, Bronik, & Fry, 1997). The main task was to remember the digits or the locations of the Xs in the order in which they were presented. However, participants also needed to simultaneously execute a secondary task, which required either a verbal response (naming the colors of the digits or Xs) or a spatial one (pointing to the color of each digit or X within a spatial array of colors).

Not surprisingly, undergraduates recalled more information than 8-year-olds, and 10-year-olds more than 8-year-olds. More interesting, at all ages, having to perform the spatial secondary task interfered to the greatest degree with recall of the spatial information and having to perform the verbal secondary task interfered most heavily with the verbal task. This finding supported the view that spatial and verbal information are represented separately in working memory. Especially interesting, at age 8 but not thereafter, having to perform a spatial secondary task also interfered with ability to recall verbal material, and having to do a verbal secondary task interfered with ability to recall spatial material. This suggests that not until age 10 do children cleanly separate verbal from spatial information in working memory.

There are also developmental changes in the executive processes that control the content and functioning of working memory, such as the ability to inhibit attention to particular sources of information when appropriate. One compelling illustration comes from a task in which children are asked to sort cards with pictures of objects (such as red boats and blue flowers) into categories based on color or shape. Children are first asked to sort the cards along one dimension (such as color), and after several trials they are asked to "switch" and sort the cards along the other dimension (shape). Three-year-olds can easily sort the cards on the basis of either color or shape, but when asked to switch dimensions, most children fail and instead continue to sort the cards on the basis of the first dimension. By 4 years of age, however, most children can successfully perform the switch (Zelazo, Frye, & Rapus, 1996). One explanation for this developmental shift is that there are age-related improvements in executive processes that control the functioning of working memory.

**Long-term memory.** Even young children are able to remember a vast assortment of experiences and facts about the world. Some of their knowledge is about specific episodes, such as their feelings when they wandered around the playground on their first day of school; this type of information is often referred to as episodic knowledge. Other knowledge is about enduring qualities of the world, such as that a nickel is worth five pennies; this type of information is often referred to as semantic knowledge. Yet other knowledge concerns procedures,
such as how to ride a bicycle; this type of information is often referred to as *procedural knowledge*. These varied types of knowledge are the contents of long-term memory.

Unlike sensory and working memory, there are no limits on either how much information can be maintained in long-term memory or how long the information can stay there. Consider an experiment on recognition of faces in high school yearbooks (Bahrick, Bahrick, & Wittlinger, 1975). People were asked 35 years after graduation to recognize which yearbook pictures were of people in their high school class and which were of people from a nearby high school. In spite of all of the time that had passed, people correctly recognized 90 percent of the pictures. Thus, the name *long-term memory* is truly a fitting one.

An interesting property of the way people store information in long-term memory is that the storage is not in all-or-none form. Rather, people store information in separable units and can retrieve some units without retrieving others. This quality has been demonstrated in adults in experiments on the tip-of-the-tongue phenomenon. When adults can almost but not quite remember a word, they often can recall several of its characteristics: its first letter, its number of syllables, a word it sounds like, and so on (Brown & McNeill, 1966). This description appears to apply to children's storage of information in long-term memory as well. For example, in trying to remember the name of a friend who had moved away, Siegler's 6-year-old daughter said, "She was from South America, she had black hair, she was just as silly as I am, why can't I remember her name?" A few minutes later, she succeeded in recalling the friend's name, Gabriella.

**Processes**

Processes are used to actively manipulate information in sensory, working, and long-term memory. Two processes that play particularly important roles in cognitive development are automatization and encoding.

*The role of automatization.* Processes vary considerably in how much attention they require. Those that require a great deal of attention are often labeled *controlled*, whereas those that require little if any attention are labeled *automatic*. The amount of attention required is influenced both by the type of information being processed and by the amount of experience the child has had processing that type of material. Some types of information inherently require less attention than others. However, even with processes that at first require a great deal of attention, practice reduces the amount that is needed.

Automatic processing is important in development in that it provides an initial basis for learning about the world. One example involves frequency information, that is, data on how often various objects and events have been encountered. People retain this information even when they are not trying to do so. Thus, we have a good sense of the relative frequency with which letters of the alphabet appear (for example, if "e" or "t" is more frequent in English), although no one tries to remember such trivia. Recall of such information is influenced neither by instructions to remember nor by practice in trying to remember it. Level of recall also is equivalent over a wide age range. Children as young as 5 are as proficient as college students at retaining frequency information (Hasher & Zacks, 1984).

Children's automatic retention of information about frequencies seems to contribute to cognitive development in many ways. When children form concepts, they must learn which features go together most frequently. For example, learning the concept "bird" requires learning that the same animals tend to fly, have feathers, have beaks, and live in trees. Likewise, in learning language, infants learn which sounds tend to occur together, and they use this information to identify words within the stream of speech (Staaf et al., 1996). More subtle learning, such as learning of sex roles, also may depend on automatic processing of frequency information. When children see a large difference in the frequency with which men and women engage in an activity, they imitate same-sex models more often than ones of the opposite sex (Perry & Bussey, 1979). Children are almost never conscious of gathering information about how often men engage in an activity and how often women do. Rather, they seem to acquire the information automatically and then base their behavior on what they have observed.

Thus, processing of frequency information appears to be automatic from early in development, perhaps from birth. Other processes, however, may change from controlled to automatic as people gain experience with them. This process is known as *automatization*.

The term "automatization" is well chosen. Once skills are learned to a sufficiently high degree, they are difficult to inhibit even when it is advantageous to do so. Learning of single-digit addition provides an example of this phenomenon, as illustrated in a study by LeFevre, Bisanz, and Mrkonjic (1988). Their experiment involved presentation of a problem such as 4 + 5 and then, a fraction of a second later, presentation of a single digit such as 9 slightly to the right of the first two numbers. The task was to say whether the number on the right was one of the addends in the problem. Thus, the answer for the above problem would be "No," because 9 was not one of the addends in 4 + 5. However, automatized knowledge of arithmetic facts would interfere with performance on this task, leading children either to say "yes" or to take longer to say "no" when the number to the right was the answer to the addition problem than when it was neither the answer to the problem nor one of the digits.

Studies of this task indicate that the easiest single-digit addition problems are automatized quite early in learning, but that it takes several years before harder ones are (LeFevre & Kulak, 1994; LeFevre, Kulak, & Bisanz, 1991; Lemaire, Barret, Fayol, & Abdi, 1994). Second graders show the interference effects associated with automatic processing only on small number problems (both addends of five or less). Third graders show the effects on both small and medium problems
(one addend of six or more), but not on large number problems (both addends of six or more). Fourth and fifth graders and adults show automatic processing on all single-digit addition problems: small, medium, and large.

As suggested by this example, automatization generally is useful, because it frees mental resources for solving other problems. For example, automatizing the addition facts would make it easier to do long multiplication problems in one's head. However, when the given problem looks like a typical problem but requires different processing, automatization can be harmful. For example, automatic activation of addition knowledge can interfere with children's performance on mathematical equivalence problems, such as $3 + 4 + 5 = 3 + \_$. Such problems resemble addition problems, but differ from them in a crucial way—the position of the equal sign. Thus, depending on the circumstance, automatization can be either harmful or helpful.

**The role of encoding.** People cannot represent all features of the environment; the world is simply too complex. Children often fail to encode important features of objects and events, sometimes because they do not know what the important features are and sometimes because they do not know how to encode them efficiently. This failure to encode critical elements can limit the effects of potentially useful experiences; when children do not take in relevant information, they cannot benefit from it.

Kaiser, McCloskey, and Poffitt (1986) provided a compelling demonstration of how inadequate encoding can hinder learning. They presented 4- to 11-year-olds and college students with a moving electric train carrying a ball on a flatcar. At a predesignated point, the ball dropped through a hole in the moving flatcar and fell several feet to the floor. The task was to predict the trajectory of the ball as it fell.

More than 70 percent of the children and a sizable minority of the college students predicted that the ball would fall straight down. After they advanced this hypothesis, the experimenter demonstrated what actually happened. (The ball moved in a curving path, going forward as well as down.) The children and the college students were faced with reconciling their predictions with the outcome they had seen. Their explanations revealed how expectations influence their encoding of what they saw. Some said that the train actually had fallen straight down but that it was released from the train later than the experimenter said it was. Others said that the train gave the ball a push forward just before it was released. Interestingly, a number of the college students who encoded the ball as having gone straight down had previously passed college physics courses that included the relevant concepts. However, this experience was insufficient to change either their expectations or their encoding of what they saw.

Encoding begins to play an important role in both developmental and individual differences in the first year of life (Colombo, 1993, 1995). Evidence for its importance comes from studies of the rate at which infants take in all of the relevant information and therefore become bored with looking at an object and look elsewhere. The length of time it takes before infants stop looking at a given object drops by more than half between ages 3 and 7 months. Recall also from Chapter 1 (p. 12) that the more rapidly 7-month-olds habituate to a repeatedly displayed object, the higher their IQs as much as 7 or 8 years later. Presumably, more intelligent infants are quicker to encode everything of interest about the picture, leading them to be the first to lose interest in it. They perk up more when the new picture is shown, because they more clearly encode the differences between it and the old one.

### Information-Processing Theories of Development

In the remainder of this chapter, we consider five types of theories of how information-processing capabilities develop: neo-Piagetian theories, psychometric theories, production-system theories, connectionist theories, and evolutionary theories. Each of these is best viewed as a family of theories, with the individual theories of each type sharing basic principles but also having unique features. The discussion of each family of theories includes identification of the shared general principles and description of one particular realization of those principles. The hope is that this discussion will convey the central features of the approach as well as a specific sense of how the approach is useful for understanding children's thinking.

All of these theories reflect the contributions of both Piaget's theory and adult information-processing approaches, as well as a number of other influences. Table 3.2 lists some of these. It also summarizes the goals of the theories and the mechanisms of development that they emphasize.

### Neo-Piagetian Theories

The goal of neo-Piagetian theories is to maintain the strengths of Piaget's approach while adding the strengths of information-processing approaches. Typically, they incorporate stages much like Piaget's with the emphasis on goals, working memory limitations, and problem-solving strategies typical of information-processing approaches. Their greatest emphasis tends to be on how the biologically based growth of working memory and automatization of processing allow children to progressively overcome processing limits. Among the most prominent neo-Piagetian theories are those of Hald (Andrews & Hald, 2002; Hald, 1993; Haldor, Wilson, & Phillips, 1998), Fischer (Fischer & Furr, 1988; Mascolo & Fischer, 1999), and Demetriou (DEMIOTI, CHRISTOU, SPANoudis, & Platsidou, 2002; Demetriou, Ekildes & Platsidou, 1993; Demetriou & Raptopoulos, 1999).

Probably the most prominent neo-Piagetian theory is that of Robbie Case. This theory can be divided into two main parts: the developmental stages themselves and the transition processes that produce progress between stages (Figure 3.2).
TABLE 3.2 Overview of Information-Processing Theories of Development

<table>
<thead>
<tr>
<th>Type of Theory</th>
<th>Representative Theorist</th>
<th>Goal of Theory</th>
<th>Main Developmental Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neo-Piagetian</td>
<td>Case</td>
<td>To unite Piagetian and information-processing theories of development.</td>
<td>Automatization, biologically based increases in working memory, and strategy construction.</td>
</tr>
<tr>
<td>Psychometric</td>
<td>Sternberg</td>
<td>To provide an information-processing analysis of the development of intelligence.</td>
<td>Strategy construction, encoding, and automatization.</td>
</tr>
<tr>
<td>Production System</td>
<td>Klahr</td>
<td>To demonstrate via computer simulation how the cognitive system modifies its own operation.</td>
<td>Generalization, based on the working of regularity detection, redundancy elimination, and the same line. Also encoding and strategy construction.</td>
</tr>
<tr>
<td>Connectionist</td>
<td>MacWhinney</td>
<td>To explain how children can learn language from the data available to them.</td>
<td>Associative competition among simple processing units. Also generalization.</td>
</tr>
<tr>
<td>Evolutionary</td>
<td>Siegler</td>
<td>To understand how the processes of variation and selection shape cognitive development.</td>
<td>Associative competition among strategies. Also strategy construction and generalization.</td>
</tr>
</tbody>
</table>

Like Piaget, Case (1985) hypothesized that children progress through four developmental stages. He characterized these stages in terms of the types of mental representations and operations children can form while they are in them. The first stage involves sensorimotor operations. Children's representations in this stage are composed of sensory input, and the actions they produce in response to these representations are physical movements. In the representational operations stage, children's representations include concrete internal images, and their actions can produce additional internal representations. In the stage of logical operations, children represent stimuli abstractly; they can act on these representations with simple transformations. In the formal operations stage, children also represent stimuli abstractly, but they are capable of performing complex transformations of the information. In each stage, children also produce representations and actions like those they produced in earlier stages.

Examples may clarify the differences in the representations that become possible in each stage. A sensorimotor operation might involve a child's seeing a frightening face (the sensory representation) and then fleeing from the room (the motor action). A representational operation might involve the child's producing a mental image of the same frightening face (the internal representation) and using the image to draw a picture of the face (the representational action). A logical operation might involve a child's realizing that two of his friends did not like each other (the abstract representation) and telling them that they could have more fun if they all were friends (the simple transformation). A formal operation might involve the child's realizing that such direct attempts at producing friendships rarely succeed (the abstract representation) and therefore leading all three into a situation in which they would need to overcome some common obstacle, thus producing friendly feelings (the complex transformation). The resemblance to Piaget's stages of development seems clear.

Case's view of the developmental sequence by which children acquire understanding of particular concepts also resembles Piaget's views. Like Piaget, Case postulates broad unities in the developmental sequence across different concepts. His views are more moderate, in that the postulated similarities of reasoning are limited to particular types of knowledge. However, even across these types of knowledge, substantial commonality in the developmental sequence is evident. In particular, Case claims that much of children's thinking is organized into central conceptual structures. A central conceptual structure is defined as "an internal network of concepts and conceptual relations which plays a central role in permitting children to think about a wide range of (but not all) situations at a new epistemic level" (Case & Griffin, 1990, p. 224). Case and his colleagues (Case, 1998; Case & Mueller, 2001; Case & Okamoto, 1996; Griffin, Case, & Sandiesen, 1992; Marinis, 1992) have focused on three main central conceptual structures, one for thinking about numbers, one for thinking about space, and one for thinking about stories. All three have a general resemblance, based on overall structural limits of the cognitive system at that age, but all also reflect the particulars of the domain to which they apply. For example, Case and Okamoto
(1996) proposed that at age 6, the central conceptual structures focus on a single dimension. In the structure dealing with numbers, this involves forming a mental number line, which allows children to perform such tasks as knowing which numbers are bigger than which other numbers. In the structure dealing with stories, the central conceptual structure at this age allows children to form mental story lines, that is, to represent the plot line of the story in terms of the sequence of events. In the structure dealing with space, 6-year-olds' thinking focuses either on the shape or location of objects, but not both. By age 8, central conceptual structures coordinating two dimensions are formed. In the domain of numbers, this allows children to coordinate two number lines, for example to understand the base 10 system dealing with numbers below 100 by coordinating understanding of 10s and 1s. In the domain of stories, it allows them to coordinate two story lines into a single plot. In the domain of space, it allows them to simultaneously represent the shapes and locations of objects.

Where Case differs most clearly from Piaget, and shows the strongest influence of the information-processing approach, is in his account of transition mechanisms. Case emphasizes working memory capacity as a determinant of cognitive growth. His claim is not that the absolute capacity of working memory increases, but rather that it functions increasingly efficiently, and that it therefore can handle more information.

How might such increases in processing efficiency occur? Case (1985) proposed that one contributor is automation. With practice, a cognitive operation that previously required all working memory resources could be accomplished more efficiently. This would free up part of the working memory capacity for other processing. It may be useful to think of this view of working memory in terms of an analogy to a car's trunk. The capacity of a car's trunk does not change as the owner acquires experience in packing luggage into it. Nonetheless, the amount of luggage that can be packed into the trunk does change. Wherewith the trunk at first might hold three suitcases, it eventually comes to hold four or five. With more efficient packing, trunk space is freed for additional cargo. Like knowledge of how to pack one's trunk, the central conceptual structures provide efficient ways of organizing goals and procedures for accomplishing the goals. Thus, they allow children to circumvent working memory limits.

Biological maturation was also assigned a role in explaining the increasing efficiency of working memory. Case (1992b) proposed that stage transitions arise from pervasive changes in electrical activity in the frontal lobes, a part of the brain particularly active in problem solving and reasoning. The specific proposal was that at the beginning of each stage, new short-distance connections develop between the frontal lobe on the left side of the brain and parts of the brain that previously had not been connected to them. During a second substage, longer-distance connections are formed within both left and right hemispheres of the brain. During a third substage, short distance connections are formed within the right hemisphere. Then the brain is ready for a new stage to begin. Patterns of electrical activity in the brain at different ages lent some support to this proposal (Thatcher, 1992).

Case and his colleagues have applied this theory to an exceptional variety of tasks. They range from scientific reasoning (Marini, 1992) to musical sight reading (Capodilupo, 1992), solving arithmetic word problems (Okamoto & Case, 1996), telling time (Case, Okamoto, Henderson, McKeough, & Bleiker, 1996; Case, Sandieson, & Dennis, 1987), handling money (Case, Okamoto, et al., 1996), drawing (Case, Stephenson, Bleiker, & Henderson, 1996; Dennis, 1992), and understanding social and emotional phenomena such as feelings, motives, and interpersonal conflict (Bruchkowsky, 1992; Case, Okamoto, et al., 1996).

Another strength of Case's theory is its usefulness for designing effective instructional techniques. Case and his colleagues have developed instructional techniques and curricula that are based on two important components of his theory: analyses of the working-memory demands of various approaches to solving problems, and analyses of the central conceptual structures that underlie core domains.

Consider Case's analysis of missing addend problems of the form 4 + ? = 7 (Case, 1985). Although the task appears simple, first graders who are taught in school find it a major obstacle. After analyzing several correct and several commonly used incorrect strategies for solving missing addend problems, Case noted that most correct strategies required more working memory capacity than 6- and 7-year-olds usually possess. However, he also noted that the simplest correct strategy and the most demanding incorrect strategy made the same memory demands. The least demanding correct strategy, according to his analysis, was to count on from the one addend given in the problem and to note the number of counts required to reach the sum. On the problem 4 + ? = 7, this simplest correct strategy would involve starting at 4 and keeping track of the number of counts needed to get from there to 7. The most demanding (and the most common) incorrect strategy was to count up first to the addend that was given and then to count on from there the number of times indicated by the sum. Illustratively, on the problem 4 + ? = 7, children would first count to 4, then count up 7 more times to 11, and finally answer that the missing addend was 11. Case reasoned that if 6-year-olds could learn the incorrect strategy, they also could learn the correct one, because the two strategies made similar demands on working memory.

The instructional strategy that Case used was straightforward. As shown in Figure 3.3, the first step was to illustrate that the equal sign (=) meant that entities on each side of the sign were equivalent. The next step (third pair of faces) was to illustrate that the plus sign (+) meant that the child should sum the entities adjacent to the sign. After the child finished working with the faces, the focus of the instruction shifted to direct consideration of problems involving numbers. In one part of this instruction, the experimenter demonstrated the incorrectness of children's existing strategy for solving missing addend problems involving numbers by having them compare the numbers on the two sides of
FIGURE 3.3 Faces used by Case (1978) to teach missing addend problems. The first pair of faces was used to demonstrate the meaning of the equal sign. The second pair was used to test whether the child could make the right-hand face equal to the left-hand face. The third pair was used to demonstrate that the whole on the right could be created from the parts on the left. The fourth pair was used to test whether the child could create a whole out of the parts. The fifth pair, like the missing addend problem, showed one part of the whole on the left and the whole on the right; the task was to fill in the other part of the whole on the left. The sixth pair included the plus sign, to make problem even more like standard missing addend problems with numbers.

the equal sign that their strategy yielded. This would allow them to see that on \(4 + ? = 7\), the 11 that their strategy yielded does not make the entities on the two sides of the equal sign equivalent. Following this, the simplest correct procedure for solving missing addend problems (the count-on strategy described in the previous paragraph) was introduced, one step at a time.

Case (1978) reported that his teaching strategy allowed 80 percent of kindergarten children to learn to correctly solve missing addend problems. This percentage represented a considerable improvement over the 10 percent of children who were able to learn such problems from the standard State of California arithmetic workbook.

More recently, Case has applied the notion of central conceptual structures to the design of curricula for a range of mathematical concepts, including rational numbers and functions (Kalchman, Moss, & Case, 2000; Moss & Case, 1999). For each domain, Case and his colleagues began by outlining the content of the central conceptual structure that underlies skilled, fluent performance within the domain, and then designed instruction that would foster the development of that central conceptual structure. For example, the central conceptual structure for rational numbers was hypothesized to incorporate children's intuitive understanding about proportions (e.g., their understanding of "half full" and a quarter full" as applied to beakers of water) and their numerical understanding of halving and doubling numbers (Moss & Case, 1999).

In Case's view, instruction should foster and extend the naturally occurring processes by which components of the central conceptual structures are coordinated and integrated. Based on this idea, Moss and Case (1999) developed an experimental curriculum about rational numbers that focused on aspects of the hypothesized central conceptual structure. For example, at the outset of the series of lessons, students were asked to use percentages to describe the fullness of various beakers of water. Eventually, percentages were linked to decimal fractions. This was achieved using large, laminated number lines with each number 1 meter apart, which were set up on the classroom floor. Students were asked to walk some percentage of the distance between two numbers (such as 75 percent), and it was explained that this value could be represented with a two-place decimal fraction (0.75 meters). Eventually, common fraction notation (1/2, 1/4, 1/8) was introduced by linking it to proportions and decimal fractions. The classroom activities were designed with the goal of strengthening and integrating aspects of the central conceptual structure for rational numbers.

Moss and Case implemented their experimental curriculum in a classroom of fourth-grade students, and they compared students' learning about decimals, fractions, and percentages with that of students in a control classroom from a demographically similar, nearby school that used a traditional curriculum. On a posttest that followed the rational number unit, students who received the experimental curriculum performed far better than students who received the control curriculum (69 percent vs. 39 percent correct), and their performance demonstrated a deeper understanding of rational numbers. The errors of students in the control group often reflected confusions of rational numbers and whole numbers; such errors were much less common in the experimental group. Students in the experimental group frequently referred to proportion concepts in justifying their solutions to problems involving decimals, fractions, and percents, and they were able to successfully solve problems that required overcoming misleading cues or generating new procedures. The experimental curriculum thus appeared to foster a generalized, flexible understanding of rational numbers.

These examples illustrate that Case's approach and, in particular, his analyses of working-memory demands and central conceptual structures are useful for applied as well as theoretical purposes. Moreover, these examples highlight the value and effectiveness of instructional techniques that are grounded in psychological theory.

Several criticisms of Case's theory have been voiced. Flavell (1984) noted that Case has not explicated the principles by which he determines how much working-memory capacity a procedure requires. As a result, it is often difficult to evaluate whether the estimates are comparable from one task to the next. Further, his ideas about the role of biological changes in producing stage changes...
are quite speculative; as yet, there is little relevant evidence available. On the other hand, Case's theory is exceptional among information-processing approaches to development in its attempt to relate basic capacities, strategies, and learning. It has yielded compelling analyses of development on many tasks and has proved practically useful as well. Also, there is a strong intuition among many researchers that improved ability to surmount memory limits does underlie much of cognitive development, though it is difficult to provide evidence that unambiguously supports the position. Thus, it seems that Case and his colleagues have taken a difficult but potentially rewarding path. To the extent that the effort succeeds, it will be a grand achievement.

**Psychometric Theories**

*Psychometric theories* are aimed at clarifying the processes measured on tests of mental abilities, such as intelligence tests. Since the beginning of the twentieth century, intelligence has been characterized by a single number, the IQ score. This practice has several drawbacks: A single number is inherently inadequate to capture a quality as rich and complex as intelligence, IQ tests may be culturally biased, and such tests do not directly measure the ability to learn and create, or the ability to apply intelligence in practical situations. IQ tests also have unique virtues, however: Scores on them are closely related to school performance at the time they are given; they predict later school performance quite accurately; and they provide a solid base from which to examine individual differences in cognitive functioning.

A number of investigators have tried to preserve these virtues while reducing or eliminating the negative qualities (Anderson, 1992; Ceci, 1990; Gardner, 1993). Probably the most prominent such theory is Robert Sternberg's *triarchic theory of intelligence*. He has applied the analysis to diverse tasks and diverse groups of children and has related his results to those yielded by traditional intelligence tests.

According to the triarchic theory (Sternberg, 1985, 1997, 1999), there are three primary aspects to human intelligence: analytical, creative, and practical. Analytical intelligence is the type of intelligence that is evaluated in most traditional tests of intelligence. It involves abilities such as analyzing, evaluating, comparing, contrasting, and criticizing. Creative intelligence comprises the abilities needed to cope with novel situations. It involves abilities such as creating, discovering, imagining, and inventing. Practical intelligence is utilized in addressing the problems that arise in everyday life, and in adapting to, shaping, and selecting environments. It involves abilities such as using and applying information.

Sternberg has argued that a common set of processes underlie analytical, creative, and practical intelligence. These processes include performance components, knowledge acquisition components, and metacomponents (Figure 3.4).

![Figure 3.4 A schematic diagram of Sternberg's theory of intelligence.](image)

**Performance components** are the processes involved in actually solving a given problem. Sternberg identified four performance components that people use to solve a great many problems: encoding, inference, mapping, and application. The way in which these performance components work can be illustrated by thinking about analogy problems. Consider the problem:

**Turkey**: Cranberry sauce :: Eggs: (1) Corn (2) Ham

The task is to decide whether corn or ham has the same relation to eggs that cranberry sauce has to turkey.

Sternberg suggested that the first step in solving this problem is to encode the terms. This step involves identifying each term's attributes—for example, noting that turkey is a kind of food, that it is a meat, that it is eaten on Thanksgiving, and so on. Next, inference is used to specify the relation between the first and second term, in this case that turkey is often eaten with cranberry sauce. Then, mapping is used to establish the relation between the first and third terms, that turkey and eggs are both foods. Finally, application involves inducing a relation between the third term and one of the possible answers that parallels the relation between the first and second terms. Here, eggs go with ham in much the same way that cranberry sauce goes with turkey.

**Knowledge acquisition components** are processes involved in learning to solve problems and acquiring relevant knowledge in the first place. Sternberg has focused in particular on three knowledge acquisition processes: selective encoding, selective combination, and selective comparison. Selective encoding involves distinguishing relevant from irrelevant information. Selective combination involves integrating information in a meaningful way. Selective comparison...
involves relating newly encoded or combined information to previously stored information.

The importance of knowledge acquisition components in solving problems can be illustrated with respect to insight problems, such as the following one: "If you have black socks and brown socks in your drawer, mixed in the ratio of 4 to 5, how many socks will you have to take out to be sure of having a pair of socks of the same color?" For this problem, skill in selective encoding is necessary to ignore the irrelevant information about the 4:5 ratio of the two colors. If you had socks of two colors, how many socks would you need to look at to be sure that two would match? When this information is present, one needs to ignore it and selectively encode only the essentials of the problem. When the irrelevant information is absent, skill in selective encoding is less important, because there is less distracting information. Not surprisingly, children are more successful on such problems when the irrelevant information is omitted, and selective encoding is not required.

Some evidence suggests that children with high IQs execute knowledge acquisition processes more effectively than other children. Consistent with this view, children with high IQs benefit less than children with average IQs from instruction that focuses on such processes. The likely reason is that children with high IQs have strong knowledge acquisition skills in the first place, so instruction has little added benefit (Davidson & Sternberg, 1984).

Metacomponents are executive processes that govern the use of the other components. The metacomponents plan how to solve problems, construct problem-solving strategies, monitor progress, and evaluate performance. They also are responsible for most aspects of developmental change. As Sternberg (1984) commented, "There can be no doubt that in the present conceptual scheme, the metacomponents form the major basis for the development of intelligence" (p. 172).

The importance of metacomponents is evident in people's transfer of knowledge from one context to another. Older children and people with greater expertise are generally better able to apply their knowledge to new problems than are younger people and people with less expertise (Campione & Brown, 1984; Gentner, Ratterman, Markman, & Kotovsky, 1995; Staszewski, 1988). Knowledge is especially important; 10-year-olds who are expert at chess more successfully solve novel chess problems than adults with little knowledge of chess but whose general memory capacities are higher (Chi, 1978). However, within a given level of knowledge, people with higher IQs generally can apply existing knowledge to acquire new knowledge more rapidly (e.g., Johnson & Mervis, 1994).

The three types of basic processes—metacomponents, performance components, and knowledge acquisition components—work together in solving problems. The metacomponents serve as a strategy construction mechanism, orchestrating the other two types of components into goal-oriented procedures. When the child already possesses sufficient understanding to solve a problem, only the metacomponents and the performance components are needed to construct a problem-solving strategy. The metacomponents select which performance components to use and the order in which to use them. The performance components do the work of actually solving the problem. If the child does not yet possess sufficient understanding to solve the problem, the knowledge acquisition components also come into play. That is, the knowledge acquisition components obtain new information relevant to solving the problem and communicate this information to the metacomponents. The metacomponents then combine the new and previous understanding to construct a problem-solving strategy.

According to Sternberg (1999), these same basic processes underlie analytical, creative, and practical intelligence; the specific processes that are applied depend on the nature of the task and the situation and on the type of thinking that is required. However, despite this fundamental similarity in terms of basic processes, traditional intelligence tests and traditional methods of classroom instruction have focused primarily on analytical intelligence, to the exclusion of creative or practical intelligence. In recent work, Sternberg has focused on designing knowledge assessments and instructional methods that address practical and creative intelligence, as well as analytical intelligence.

Tests that assess all three types of thinking have proven to be more effective at predicting intellectual outcomes than traditional IQ tests, which focus on analytical abilities alone. One study of this issue involved gifted high school students who were selected for a college-level, summer psychology course. Students completed tests of analytical, creative, and practical intelligence, and their scores on these tests were used to predict their grades in the course. Grades were better predicted based on the combination of abilities than based on analytical ability alone (Sternberg, Ferrari, Clankinbeard, & Grigorenko, 1996).

This study also involved an instructional component. All of the students used the same psychology textbook and listened to the same lectures. However, students were assigned to discussion sections that differed in their focus on analytical, creative, practical, or memory activities. Sternberg and colleagues found that students who were placed in discussion sections that matched their strengths (e.g., students with creative strengths in the section that focused on creative activities) performed better in the course than students who were placed in discussion sections that did not match their strengths. Thus, students perform better when the instructional conditions match the way they think best.

Other studies have shown that instruction that emphasizes all three types of intelligence is more beneficial for students than either conventional instruction or instruction that focuses on critical-thinking skills. This pattern has been documented in several different participant populations and with different types of course content, including third-grade students learning a social studies unit; eighth-grade students learning a psychology unit; inner-city fifth-grade students learning reading skills; low-SES middle school students in a summer reading enrichment program; and high school students in multiple subject areas (Grigorenko, Jarvin, & Sternberg, 2002; Sternberg, Torff, & Grigorenko, 1998). It
It seems clear that instruction that encourages creative and practical thinking as well as analytical thinking can be beneficial for many students.

How should Sternberg's theory be evaluated? Three important weaknesses can be noted. One is that the theory summarizes more than it predicts. It is not clear what types of evidence would be inconsistent with the approach. A second weakness has to do with the specification of basic processes. The functioning of these processes has been spelled out primarily for tasks that involve analytical thinking; it is less clear how they function in tasks that require creative and practical thinking. A third weakness involves the role of metacomponents in the organization of the system. The metacomponents are crucial parts of the overall theory, but their workings remain somewhat mysterious.

On the other hand, the theory is exceptional in the breadth of phenomena and of populations to which it has proven applicable. It encompasses a large number of intuitively important aspects of development and organizes them in an easy-to-grasp way. It provides a plausible outline of how a strategy-construction mechanism would operate. It has yielded important practical applications both for the assessment of intelligence and for instruction. In short, it constitutes a useful framework within which to view development, and it also has yielded substantial practical benefits.

**Production System Theories**

Perhaps the most difficult challenge for theories of cognitive development has been to explain how development occurs. Piaget and many others have tried to generate such explanations, but they have not been entirely successful. Consider the following evaluation:

> For 40 years now we have had assimilation and accommodation, the mysterious and shadowy forces of equilibration, the Batman and Robin of the developmental processes. What are they? How do they do their thing? Why is it after all this time, we know no more about them than when they first sprang on the scene? What we need is a way to get beyond vague verbal statements of the nature of the developmental process. (Klahr, 1982, p. 80)

One promising effort to provide more precise and satisfying explanations of change has been to model development through production systems (Klahr & MacWhinney, 1998). These are a class of computer-simulation languages that have proved useful for modeling cognitive development. Each production is a kind of if-then rule that indicates what the system would do in a particular situation. Together, the productions indicate what the system would do under a wide range of circumstances. The key properties of production systems are the following:

1. The basic organization consists of two interacting structures: a production memory, which is the system's enduring knowledge, and a working memory, which is the system's representation of the current situation.

2. The production memory includes a large number of specific productions, each of which includes a condition side and an action side.

3. The condition side of each production specifies the circumstances under which the production is applicable. The action side specifies the actions that are taken when these conditions are met. Such actions include both activities in the external world and manipulations of symbols in working memory.

4. The contents of working memory are constantly changing, because they reflect constantly changing situations. Information enters working memory both through perception of events in the external world and through taking the actions indicated by the action side of productions.

5. Thinking occurs through a cycle of a) information being present in working memory, b) the information matching the conditions of one or more productions, c) this match resulting in the actions on the action side of those productions being taken, d) the actions placing new information in working memory, thus starting the cycle anew.

6. Learning occurs through a process of self-modification, in which new productions are created and existing productions modified as a result of previous experience.

The basic organization of production systems is diagrammed in Figure 3.5.

An example of a simple production system that generates correct performance on Piaget's number conservation problem is shown in Table 3.3. The

---

**FIGURE 3.5 The hierarchical organization of production systems.**

---

---
**TABLE 3.3 A Simple Production System for Number Conservation**

P1: If you are asked about the numerical relation between two collections and you do not have a goal of stating the relation, then set a goal of stating the relation.

P2: If you have a goal of stating the numerical relation between two collections and you know the relation, then state the relation.

P3: If you have a goal of stating the numerical relation between two collections, and the collections had the same number of objects before a transformation and the transformation did not involve adding or subtracting objects, then the collections still have the same number of objects.

Initial Working Memory (WM1): Rows had same number of objects before, one row then was spread, nothing added or subtracted, question is whether rows have same number of objects now.

P1 fires.

WM2: Goal is to state whether rows have same number of objects now, rows had same number of objects before, one row then was spread, nothing added or subtracted, question is whether rows have same number of objects now.

P3 fires.

WM3: Goal is to state whether rows have same number of objects now, rows have the same number of objects, rows had same number of objects before, one row then was spread, nothing added or subtracted, question is whether rows have same number of objects now.

P2 fires.

System answers: "The rows have the same number of objects."

Source: Adapted from Klahr & Wallace, 1976.

---

**TABLE 3.4 A Portion of a Child's Time Line**

(Previous Processing Episodes)

<table>
<thead>
<tr>
<th>Event</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>87456</td>
<td>Cookies on table.</td>
</tr>
<tr>
<td>87457</td>
<td>I subitized.</td>
</tr>
<tr>
<td>87458</td>
<td>There were three.</td>
</tr>
<tr>
<td>87459</td>
<td>I heard a bird.</td>
</tr>
<tr>
<td>87460</td>
<td>I picked up the cookies.</td>
</tr>
<tr>
<td>87461</td>
<td>I subitized the cookies.</td>
</tr>
<tr>
<td>87462</td>
<td>There were three again.</td>
</tr>
</tbody>
</table>

---

that the rows have the same number of objects. With this information, P2 can fire and the system states the correct answer.

Many researchers have utilized production system models as a tool for studying development (e.g., Jones, Ritter, & Wood, 2000; Klahr, Langley, & Neches, 1987; Young & O'Shea, 1981). One especially prominent advocate of production systems as a tool for explaining development is David Klahr. The key developmental mechanism in Klahr's production system theory is generalization. Number conservation provides a convenient context for explaining how his theory works.

Klahr and Wallace (1976) divided the process of generalization into three components: the time line, regularity detection, and redundancy elimination. The time line contains the data on which generalizations are based. It is a record of all the situations the system has ever encountered, the responses produced in those situations, the outcomes of the actions, and the new situations that arose. Table 3.4 illustrates the type of information that might be included in the time line's record of a single event. A child saw a group of cookies and noticed that there were three. This realization was made possible by subitizing (a process by which both children and adults can rapidly perceive the number of objects in sets ranging from one to four objects). Next, the child transformed the spatial position of the cookies by picking them up in his hand. Finally, the child again subitized the collection of cookies and found that there still were three.

Such detailed records of situations, responses, and outcomes might at first seem unnecessary. Why remember so much about each experience? In fact, the information could be invaluable. In many situations, children cannot know beforehand what will turn out to be relevant. If they retain detailed information that may or may not be relevant, they later may be able to draw unanticipated generalizations. If they retain only what they already know to be relevant, however, they will miss much relevant information.
Is it realistic to think that children have a memory record similar to a time line? Observing the level of detail with which they remember certain information suggests that it is. Almost all parents have anecdotes to this effect. One of Siegler’s concerns a vacation on which he and his wife and their almost-2-year-old son were staying in a motel. They wanted to go to dinner but could not find the room key. After 10 minutes of searching, the father finally listened to his son long enough to understand that he was saying: “Under phone.” The father knew immediately that his son was right. He had put the key there (for reasons he no longer remembers). It seems likely that if the child remembered this relatively inconsequential detail, he probably was remembering many other details as well. Hasher and Zacks’s (1984) ideas about automatic processing of frequency information and of several other aspects of experience, such as spatial locations and time of occurrence, suggest the types of content that might be entered into the time line. Thus, Klahr and Wallace’s contention that children retain a detailed ledger of their experiences seems quite plausible.

The second key process, regularity detection, operates on the contents of the time line to produce generalizations about experience. This is accomplished by the system’s noting places in the time line where many features are similar and where the same outcome occurs despite variations in one or more features. In number conservation, regularity detection could produce at least three types of generalizations. One would involve generalizing over different objects. Regardless of whether two checkers, two coins, two dolls, or two cookies were spread, there still would be two objects. Children also could generalize over equivalent transformations. Spreading, compressing, piling up, and putting in a circle all preserve the initial number of objects.

The third process in Klahr and Wallace’s model, redundancy elimination, accomplishes a different type of generalization. It improves efficiency by identifying processing steps that are unnecessary, thus reaching the generalization that a less complex sequence can achieve the same goal. In the number conservation example, children eventually would note that it is unnecessary to subitize again after picking up the cookies. Since there were three cookies before, and since picking up objects never affects how many there are, the number still must be the same. Klahr and Wallace hypothesized that the information-processing system eliminates redundancy by examining procedures within the time line and checking if the same outcome always occurs even if one or more steps are deleted. If so, the simpler procedure is substituted for the more complex one.

What does the information-processing system have time to detect regularities and to eliminate redundancies? Klahr and Wallace (1976) advanced one intriguing possibility: Perhaps children do it in their sleep. Other possibilities are that moments of quiet play, relaxation, or daydreaming are when children accomplish these functions.

Klahr and Wallace’s approach, unlike stage theories, implies that different children develop skills in different orders. In the cognitive system’s attempts at self-modification, there is no reason why one type of regularity always should be detected before another type. Children learning about number conservation either could first detect that it does not matter if the rows of objects contain cookies or checkers or could first detect that it does not matter if the row of cookies is shortened or lengthened. Thus, there is less of a lock-step feel to the model than there is to stage approaches.

Another implication of Klahr and Wallace’s theory relates to the idea of encoding. The way in which information is encoded in the time line shapes the learning that can later occur. Suppose, for example, that in a conservation of liquid quantity experiment, a child encodes only the heights of the water in the glasses. Such a child would not be able to detect the regular relation between increments in the height of water and decrements in its cross-sectional area. The information about cross-sectional area simply would not be available in the time line.

Klahr has been in the forefront of investigators arguing for greater use of computer simulation as a tool for modeling development. He has noted that such simulations allow more explicit and precise models of how development occurs than would otherwise be possible (Klahr, 1989, 1992; Klahr & MacWhinney, 1998). Consistent with this stance, Simon and Klahr (1995) formulated a self-modifying production system that illustrated how children could come to understand conservation. At the outset, the model could not solve the number conservation problems it was presented, but through experience trying to solve them, it figured out how to do so. Of special interest, Simon and Klahr generated two versions of the model, one corresponding to 3-year-olds and one to 4-year-olds. Both models were able to learn when given relatively extensive experience with the problems, but only the model of 4-year-olds learned from limited experience with them. These data corresponded to the results obtained with real 3- and 4-year-olds who had been presented with these experiences by Gelman (1982).

The models of the younger and older children suggested hypotheses concerning why 3- and 4-year-olds showed the patterns of learning that they did. Both models contained learning mechanisms that allowed them to learn from the more extensive experience. However, two differences between them resulted in the model of 4-year-olds, but not the model of 3-year-olds, learning from the limited experience. The model of 4-year-olds more clearly remembered the relation between the sets before the transformation, and it was more likely to check whether the differences between the lengths of the rows after the transformation corresponded to a difference in numbers of objects. These differences in the models were consistent with what is known generally about 3- and 4-year-olds. The 4-year-olds are more likely to use counting to check whether their perceptions regarding numbers of objects are correct (Sophon, 1987), and they also usually remember more about past states (Schneider & Bjorklund, 1998). Thus, the differences between the models of 3- and 4-year-olds were both consistent with past observations of these age groups and suggested hypotheses regarding why the two age groups might learn as they did in this particular context.

Not everyone shares Klahr’s enthusiasm for computer-simulation models, though. Critics note that people are not computers and that unlike computers,
people develop. This leads them to the conclusion that development cannot be
modeled appropriately on a computer (Bellin, 1983; Liben, 1987).

As Klahr (1989) pointed out, however, ideas about development are
embodied in the computer program, not the computer on which the program runs. The
computer is simply the device used to test whether these ideas account for
the known phenomena. To illustrate the point, Klahr noted that computer simulations
of cognitive development do not imply that children are computers any more than
computer simulations of hurricanes imply that the atmosphere is a computer.

Several limitations of Klahr's theory should be mentioned. Although he
has often proclaimed the virtues of self-modifying production systems, neither
he nor other investigators interested in children's thinking have yet written
many of them. In addition, such self-modifying production systems thus far
have been more useful for explaining previous findings than for generating new
ones. On the other hand, these shortcomings do not detract from the potential
of self-modifying production systems as models of development. In addition,
Klahr's explanation of generalization in terms of the time line, regularity
detection, and redundancy elimination is more precise and explicit than almost all
other mechanisms of cognitive development that have been proposed. These are
important virtues and may foreshadow additional breakthroughs.

**Connectionist Theories**

One especially "hot" approach to thinking about cognitive development (and to
thinking about cognition in general) is connectionism. Like production systems,
connectionist theories are computer simulations of how thinking occurs. Much
of the reason for the popularity of connectionist models is their general resemblance
to the workings of the brain. This makes the approach a promising candidate for
modeling how thinking is achieved within the brain. The models have
several key characteristics (Plunkett, 1996):

1. They are made up of large numbers of simple processing units, akin to neurons in
   the brain.
2. The processing units are organized into two or more hierarchically-organized layers
   (Figure 3.6). Typically, these include an input layer, whose processing units encode
   the initial representation of the situation; one or more hidden layers, whose units
   combine information from the input units; and an output layer, whose units generate
   the system's response to the situation.
3. The individual processing units are connected to other processing units in different
   layers (and sometimes within the same layer as well). The strength of each
   connection varies with the system's experience and is crucial in determining the
   processing that is done.
4. As in the brain, a given processing unit fires when the amount of activation it is
   receiving from all of the other processing units that are connected to it exceeds a
   threshold. The amount of activation that a unit receives from each unit connected to
   it is determined by the degree of activation of the processing unit that is sending
   the activation and the strength of the connection between the units.

![Figure 3.6](http://example.com/connectionist-model.png)

FIGURE 3.6 The connectionist model of MacWhinney et al. (1989) of how children
learn the German system of articles. Note that at the top (input) level, the
model encodes five semantic features of the noun (corresponding to the "5" at
the level just below the top on the extreme left), presence or absence of 11
phonological features at as many as 13 locations in the word (also represented
at the level just below the top), and 17 explicit case cues, which indicate the
function of the noun within the sentence. These input-level units transfer activa-
tion to hidden units at the next two levels below, and eventually to the six
output units, which correspond to the six articles that accompany nouns in
German. The article that goes with the most activated output unit is advanced as
the response. Reprinted from MacWhinney, B., Leinbach, J., Tabakan, R., & Mc
Donald, J., Language learning: Cues or rules? Journal of Memory and Language,
28, 255-277. Copyright © 1989, with permission from Elsevier.)
5. As in the brain, the activity of the many simple processing units occurs in parallel (simultaneously).

6. Knowledge is represented through the strengths of connections among all of the units in the system. There is no single location that corresponds to a particular piece of knowledge; rather the knowledge is distributed over all of the units and their interconnections. Because of this, and because processing occurs over many units in parallel, these systems are often known as parallel distributed processing (PDP) systems.

7. Learning occurs through the system receiving input, generating a response, observing the discrepancy between it and the correct answer, and adjusting the strengths of connections among the processing units in ways that would have led to a better answer. The adjustments include strengthening some connections and weakening others. Through this process, the system implicitly learns the rules underlying correct responses to the problem, although there is no single place in which the rule is represented.

8. Generalization of the system's knowledge is based on similarity of new situations to ones the system has encountered previously. When the same types of implicit rules apply to new problems, connectionist systems are very effective in generalizing previous experience to them.

A number of researchers have advocated the use of connectionist models as a tool for understanding development: McClelland (1995), Shultz (2003), Plunkett (1996), and Marchman (1992) among them. One particularly impressive connectionist model, and one that illustrates the strength of the approach for modeling development, is that of MacWhinney, Leinbach, Taraban, and McDonald (1989).

MacWhinney et al.'s model depicted German children's learning of their language's system of definite articles. These definite articles are the multiple terms that in German serve the function that the single word the serves in English. The task was of interest precisely because the German article system is so difficult. Which article should be used to modify a given noun depends on the gender of the noun being modified (masculine, feminine, or neuter), its number (singular or plural), and its role within the sentence (subject, possessor, direct object, prepositional object, or indirect object). To make matters worse, assignment of nouns to gender categories is often nonintuitive. For example, the word for fork is feminine, the word for spoon is masculine, and the word for knife is neuter. The relations are so complex that they seem almost impossible to learn. However, MacWhinney et al. built a connectionist model that showed how children could learn them.

The MacWhinney et al. model, like most connectionist models, involves an input layer, several hidden layers, and an output layer (Figure 3.6). Each of these layers contains a number of discrete units. For example, in the MacWhinney et al. model, the 35 units within the input layer represent features of the particular noun that the article modifies, specifically aspects of the noun's sound, meaning, and context. Each of the hidden layers includes units that represent combinations of these input-layer features. The six output units represent the six articles in German that correspond to the in English (der, die, das, dem, den, and des).

As just noted, a central feature of such connectionist models is the very large number of connections among processing units. In the MacWhinney et al. model, each input-layer unit is connected to first-level hidden units; each first-level hidden unit is connected to second-level hidden units; and each second-level hidden unit is connected to each of the six output units. Learning occurs through a cycle of the system (1) receiving initial input (in this case, a noun in a certain context); (2) projecting on the basis of the strengths of its various connections (which reflect past experience) what output to produce; (3) advancing that response; and (4) adjusting the strengths of connections between units so that connections that suggested the correct answer are strengthened and connections that suggested the wrong answer are weakened.

MacWhinney et al. tested this system's ability to master the German article system by repeatedly presenting the system 102 common German nouns. The model needed to choose which article to use with each noun in the particular context—that is, in the context of wanting to express a particular meaning with particular words. After it did this, the correct answer was presented, and the model adjusted connection strengths so as to optimize its accuracy in the future.

Following experience with this training set, the MacWhinney et al. model chose the correct article for more than 90 percent of the nouns in the original set. This could not be attributed simply to rote learning of which article accompanied each noun. When the model was presented with a previously encountered noun in a novel context, it chose the correct article on more than 90 percent of trials, despite the noun's often taking a different article in the new context than it had in the previous ones. The model also proved able to generalize to novel nouns; even when it had never encountered the particular term, it could use the term's sound and meaning to make educated guesses as to what article would accompany it.

The model's learning paralleled children's learning in a number of ways. Early in the learning process, the model, like children whose first language is German, tended to overuse the articles that accompany feminine nouns. The reason appeared to be that this form of the article is used most often within the language. Further, the same article-noun combinations that are the most difficult for German children to learn were the most difficult for the model to learn as well. The particular errors made by the model also resembled those of children.

How was it possible for the model to produce systematic behavior without learning explicit "rules" (such as "masculine nouns in dative case take the article dem")? The answer is that systematic behavior emerges from the operation of a simple mechanism (MacWhinney, 1995). Namely, the strengths of the connections between input, hidden, and output units were adjusted repeatedly to reflect the frequency with which particular combinations of noun features were associated with each article. Eventually, the pattern of connection strengths captured complex patterns of multiple, interacting cues, without the need for formal rules to specify those combinations of cues. The fact that the model could learn to apply articles correctly without learning explicit "rules" suggests that children acquiring German may also learn the system of definite articles without acquiring rules.
In constructing their model, MacWhinney and colleagues specified the number of units in each layer. The number of units in the input layer was chosen on the basis of the number of features of the input tokens that they wished to represent, and the number of units in the output layer was chosen on the basis of the number of different articles that exist in the German language. The number of hidden units was also set in advance by the researchers. As the model learns, the hidden units come to represent systematic patterns that occur among the input units, such as combinations of input features. The number of hidden units was selected by estimating how many such patterns might be important in learning the input-output relations in the domain. Models that contain different numbers of hidden units may display different patterns of learning (Quinn & Johnson, 1997).

Some recent connectionist models utilize a learning mechanism that allows the models to recruit new hidden units into the network when they are needed. In the course of learning, when the model reaches a plateau where its performance is no longer improving, it recruits a new hidden unit, and this enables new gains in learning. The recruitment of new hidden units is thought to be analogous to forming new synapses in the brain as a result of learning. A recent model of children’s acquisition of conservation of number provides a good illustration of this type of model (Shultz, 1998).

Shultz’s model incorporates 13 input units that code several types of information: (1) information about the length and density of each row, (2) information about which row is transformed, and (3) information about the nature of the transformation (addition of an item, subtraction of an item, compression of the row, or elongation of the row). The model also incorporates two output units. When the two output units have the same level of activation, the model “judges” that the rows are equal in number. When one of the output units has a greater level of activation than the other, the model “judges” that the row corresponding to that output unit has the greater number.

The model was trained on a set of 420 different conservation problems, which varied in terms of the length, density, and number of objects in the rows prior to the transformation, and the nature of the transformation (addition, subtraction, compression, or elongation). During training, the model was presented with conservation problems, and after each problem, the network connection strengths were adjusted using a mathematical algorithm to reduce errors in the output. When these connection strengths adjustments failed to improve performance, the model recruited a new hidden unit. In general, the model displayed large improvements in performance immediately after new hidden units were recruited.

After the model successfully learned to solve the conservation problems in the training set, Shultz presented it with a set of 100 items that were not part of the training set. The model successfully generalized its learning, succeeding on 95 percent of the new items. The model’s learning also paralleled children’s learning in some important ways. In longitudinal studies, children tend to show sudden jumps in conservation performance; the model displayed similar sudden improvements. Children tend to succeed on conservation problems that involve small numbers before they succeed on conservation problems that involve large numbers. The model showed this exact pattern. Children also tend to choose the longer of the two rows as having more items than the shorter row (regardless of density); the model did so as well.

Connectionist models have successfully depicted a number of other developmental acquisitions as well. These include object permanence (Munakata, 1998; Munakata, McClelland, Johnson, & Siegler, 1997), understanding of time-speed-distance problems (Buckingham & Shultz, 2000; Shultz, Schmidt, Buckingham, & Mareschal, 1995), early reading acquisition (Plaut, McClelland, Seidenberg, & Patterson, 1995), second language learning (MacWhinney, 1996), category learning (Mareschal, French, & Quinn, 2000; Quinn & Johnson, 2000), and acquisition of word meanings and grammatical understanding (Elman, 1993; MacWhinney & Chang, 1995; Marchman, 1992; Plunkett & Sinha, 1992; Shultz & Bale, 2001).

As with all theories, connectionist approaches are open to criticism. One frequent criticism is that their claim to be “brain style cognition” is overstated. Nothing within them corresponds to the chemical activity that is crucial to brain functioning, and the functioning of their simple processing units bears only an abstract similarity to the functioning of neurons. Another limitation is that connectionist networks learn extremely slowly, and they require many more exposures to learn than human beings do. Although models that recruit new hidden units sometimes show sudden gains in performance, they do not show the kind of sudden insight that people fairly often do (Raijmakers, van Koten, & Molemaar, 1996). A third limitation, related to the second one, is that they do not learn the symbolic rules, such as mathematical formulas, that people do, and they may not be able to learn certain aspects of grammar (Pinkser & Prince, 1988).

On the other hand, connectionist models have proven useful for modeling the many developments that do not depend on acquisition of explicit rules. Although the operation of such systems clearly differs from that of the brain, it more closely resembles it than do other computer simulation approaches. Connectionist models have proven especially useful for modeling domains such as perception and language, in which numerous, partially valid sources of information must be integrated to produce successful performance. Given these advantages, it is not surprising that the popularity of connectionist modeling of development is growing rapidly.

**Theories of Cognitive Evolution**

One of the most profound intellectual contributions of all time is Darwin’s theory of evolution. Within evolutionary theory, competition among species is a basic aspect of existence. Species originate and change through two main processes: variation and selection. Genetic combination and mutation produce
variation; survival of offspring is the basis of selection. Together, these processes have produced our planet's ever-changing mosaic of living things.

As in the biological context, competition seems to be a basic feature of cognition. Rather than species competing, however, the competitions are among ideas. The main challenges for evolutionary theories of cognitive development are to describe the competing entities within the human cognitive system, to describe how the competition among these entities leads to adaptive outcomes, and to identify the mechanisms that produce cognitive variation and selection.

A number of current models of cognitive development are based on analogies between the functions that must be accomplished to produce evolutionary and developmental change (Changeux & Dehaene, 1989; Edelman, 1987; Geary & Bjorklund, 2000; Johnson & Gilmore, 1996). Here we use Siegler's (1996, 2000) overlapping waves approach to illustrate the way in which the analogy to biological evolution can contribute to understanding of development.

The basic assumptions of this approach are that at any one time, children have a variety of ways of thinking about most topics; that these varied ways of thinking compete with each other for use; and that the more advanced ways of thinking gradually become increasingly prevalent. These assumptions are illustrated in Figure 3.7. At any given time, several ways of thinking (the strategies in the figure) are present in a child's thinking. (Strategies are procedures aimed at meeting particular goals.) These strategies compete with each other, and with experience, some become more frequent, some become less frequent, and some first become more frequent and later less frequent. Further, new strategies are introduced and old strategies stop being used. This overlapping waves model seems more in accord with what is known about cognitive development than do depictions that show children suddenly moving from one approach to another.

Siegler and colleagues have pursued this evolutionary model within a variety of areas: arithmetic, time telling, reading, spelling, tool use, problem solving, and memory tasks, among them (Chen & Siegler, 2000; Jansen & van der Maas, 2002; Rittle-Johnson & Siegler, 1999; Siegler, 1996; Siegler & Stern, 1998; Siegler & Svetina, 2002). In each of these areas, the findings indicate that competition leads to adaptive consequences, and that basic strategy choice and discovery mechanisms produce the adaptation. The findings can be illustrated in the context of young children's learning of simple addition.

First consider the competing entities. Even 5-year-olds use a variety of strategies to solve basic addition problems such as 3 + 5. Sometimes they count from one; this typically involves putting up fingers on one hand to represent the first addend, putting up fingers on the other hand to represent the second addend, and then counting the raised fingers on both hands. Other times, they put up fingers but recognize the number of fingers that are up without counting. Yet other times, they retrieve an answer from memory. Some children also know another strategy, the count-on strategy. Children using this strategy choose the larger of the two addends and count on from that point the number of times indicated by the smaller addend. For example, on 3 + 9, children might think to themselves, "9, 10, 11, 12."

It is not the case that some 5-year-olds use one of these strategies and some use another. Rather, almost all children use several different strategies. In addition, on arithmetic, spelling, time telling, memory recall, analogical reasoning, tool use, and many other tasks, the majority of children have been found to use multiple strategies. Even on individual problems, the outcomes of the competition vary, so that the same child will choose one strategy one day and a different one the next (Siegler, 1987a).

Children's choices among these strategies are adaptive in several different ways. One sense in which their choices are adaptive is that they use retrieval, the fastest strategy, predominantly on simple problems where it can yield accurate performance, and they use more time-consuming and effortful strategies on more difficult problems, where such strategies are necessary for accurate performance (Siegler, 1986).

Children also choose adaptively among strategies other than retrieval. In particular, they tend to use each strategy most often on problems where it works especially well compared to alternative approaches. In evolutionary terms, strategies find their niches. For example, the count-on strategy is used most often on problems such as 2 + 9, where the smaller addend is quite small and the difference between addends is large. On such problems, counting on is both easy to do and effective relative to alternative procedures such as counting from one (Siegler, 1987b).

Changes over time in strategy use also are adaptive. For example, in simple addition, children increasingly use the most efficient strategies, such as retrieval and counting on, and decrease their use of less efficient strategies, such as guessing and counting from one. They also acquire new strategies, such as decomposition (such as solving 3 + 9 by thinking "3 + 10 = 13, 9 is 1 less than 10, so 3 + 9 = 12").
strategy choices on 9 + 1. The simulation gradually learns that it is easier to solve this problem by counting from the larger addend than by counting from one. It requires far fewer counts to say “9, 10” than “1, 2, 3, 4, 5, 6, 7, 8, 9, 10.” This lesser amount of counting results in fewer errors and shorter solution times, which in turn leads to more frequent future choices of the counting-on strategy. The simulation uses this experience and similar experience with other problems to draw the generalization that counting on works better than counting from one on related problems, such as 9 + 2 and 8 + 1, and it uses the knowledge to generalize appropriately to unfamiliar problems.

As children increasingly choose strategies that correctly solve problems, they also increasingly associate the correct answers with the problems. For example, 9 + 1 becomes strongly associated with 10. This association allows them to retrieve 10 as the answer to the problem. Retrieving an answer is even faster than counting “9, 10” and is just as accurate. Thus, the very success of the count-on strategy in producing correct answers leads to its own obsolescence, because it makes accurate retrieval possible.

The evolutionary perspective raises the further issue of the source of strategic variation. In particular, how are new strategies acquired? Sometimes children are taught a new strategy or imitate another person who is using it. However, the most interesting case is strategy discovery, in which children invent a strategy for themselves.

How do children discover new strategies? To find out, Siegler and Jenkins (1989) examined 4- and 5-year-olds’ discovery of the counting-on strategy. Recall that this strategy involves solving problems such as 2 + 9 by thinking, “9, 10, 11.” Children in the Siegler and Jenkins experiment knew how to add by counting from one but did not yet know how to do so by counting on from the larger addend. The children practiced solving addition problems three times per week for 11 weeks. Because even young children can accurately report immediately after an addition problem how they solved the problem (Siegler, 1987b), it was possible to identify the exact trial on which each child first used the new strategy. This allowed Siegler and Jenkins to examine what led up to the discovery, and what the experience of discovering a new strategy was like for a child.

Almost all of the children discovered the new strategy during the course of the experiment. The time that they took to make the discovery varied widely; the first discovery came in the second session, whereas the last one did not come until the thirtieth session. The quality of the discoveries also varied widely. Some discoveries showed a great deal of insight, as exemplified by “Lauren’s” protocol:

E: How much is 6 + 3?
L: (Long pause) Nine.
E: OK, how did you know that?
L: I think I said... I think I said... oops, um... I think he said... 8 was 1 and... um... I mean 7 was 1, 8 was 2, 9 was 3.
children's strategy discovery process may work in the same way as the discovery process implemented in the model.

In some cases, children discover new strategies without being consciously aware of their discoveries. Siegler and Stern (1998) investigated this issue using inversion problems, which are arithmetic problems of the form \( a + b - b = \) (such as \( 18 + 5 - 5 = \)). Children sometimes use a shortcut strategy to solve these problems: they recognize that since the same number is added and then subtracted, that number can be ignored, and the other number is the solution. At other times, children use a more labor-intensive computation strategy; they compute the solution in two steps (\( 18 + 5 = 23; 23 - 5 = 18 \)). Not surprisingly, the computation strategy takes much longer to execute than does the shortcut strategy. In Siegler and Stern's study of second grade students, when children used computation, they solved the problems in an average of 16 seconds, and when they used the shortcut strategy, they solved the problems in an average of 2.5 seconds.

Most interesting, children sometimes solved a problem very quickly (in fewer then 4 seconds), suggesting that they used the shortcut strategy, but claimed to have used computation when asked how they solved the problem. They appeared to use the shortcut strategy unconsciously before they could report that they had used it! Siegler and Stern termed this phenomenon the unconscious shortcut strategy, and they found that it was especially prevalent among children who were given problem sets that consisted solely of inversion problems, rather than sets with problems of the form \( a + b - c = \). In the group that received only inversion problems, 14 of the 16 children (88 percent) used the unconscious shortcut strategy before they used the (conscious) shortcut strategy. It was not the case that verbal skills prevented children from expressing the shortcut strategy—almost all of the children explicitly stated the shortcut strategy on a later trial. Instead, the findings suggest that children used the shortcut strategy without being aware of it, before they could consciously recognize that they had used it.

These findings may shed some light on Lauren's difficulties in articulating her discovery of the counting-on strategy in the protocol presented above. At least in some cases, the process of strategy discovery appears to involve cognitive processes that are not easily verbalized. Other researchers have also reported evidence consistent with this view. For example, Goldin-Meadow and colleagues have demonstrated that children often express new problem-solving strategies in gestures before they express them in speech (Alibali & Goldin-Meadow, 1993; Church & Goldin-Meadow, 1986; Goldin-Meadow, 2001; Perry, Church, & Goldin-Meadow, 1988). These researchers have argued that, at least in some cases, children's emerging knowledge is implicit and not accessible to verbal report. Eventually, this knowledge is re-represented in a more explicit format, which is accessible to verbalization. Along similar lines, Dixon and Moore (1996) have argued that solvers must have intuitive understanding of a problem domain before they can generate problem-solving strategies within that domain.
What are the main limitations of Siegler’s theory? One problem is that the theory seems most applicable to domains in which children use clearly defined strategies; its applicability to areas in which strategies are less well defined remains to be demonstrated. Another is that it has little to say about how the social world influences cognitive development. Still, it would be disingenuous to be pessimistic about it. The basic observation that cognitive development resembles biological evolution is beginning to emerge in many areas: perceptual development (Johnson & Karmiloff-Smith, 1992), language development (MacWhinney & Chang, 1995), motor development (Thelen & Fisher, 1989), and analogical reasoning (Gentner, 1989) among them. If the approach proves half as useful in understanding cognitive development as it has in understanding biological evolution, the effort to apply the idea will be well worthwhile.

Summary

Information-processing theories of development have several distinguishing characteristics. Their basic assumption is that thinking is information processing. They emphasize precise analysis of change mechanisms. They focus on the strategies that children devise to surmount the challenges posed by the environment and by their own limited processing capacity and knowledge.

Within information-processing approaches, cognition is viewed as reflecting both structure and process. Structure refers to relatively fixed aspects of the information-processing system, process to relatively variable and changeable ones. Among the most critical structures are sensory, working, and long-term memory. Sensory memory is devoted to holding a relatively large amount of unanalyzed information for about a second after the information is encountered. Working memory involves the information in the current situation and in long-term memory that is receiving attention at any given time. Without continuous attention, information is lost from working memory within 15 to 30 seconds. Long-term memory involves our enduring knowledge of procedures, facts, and specific events. It appears to be of unlimited capacity, and information remains in it indefinitely.

In contrast to this relatively small number of structures, each of which influences thinking in almost all situations, a much larger group of processes contributes in more delimited situations. These processes vary greatly with the particular circumstances, thus giving human cognition much of its flexibility. The same situation also elicits different processes in different people, depending on their past experience and abilities. Rules, concepts, and strategies are among the types of processes that people most often use.

Several information-processing theories of development have been formulated to make understandable how creatures as helpless and ignorant as infants eventually attain the power and flexibility of the adult information-processing system. Neo-Piagetian theories are aimed at uniting Piagetian and information-processing theories. Case’s approach is a particularly influential example. It posits a series of stages much like Piaget’s and a set of central conceptual structures that organize thinking in domains such as number, space, and stories. It also suggests that limited working memory capacity is a major obstacle to cognitive growth. By automatizing their processing, through biological maturation, and through acquisition of more advanced central conceptual structures, children become able to perform increasingly difficult cognitive feats.

Psychometric theories are intended to reveal the processes underlying the individual differences that appear on intelligence tests. Sternberg’s triarchic theory of intelligence illustrates how information-processing ideas can be used to pursue this goal. The theory holds that there are three primary aspects to human intelligence: analytical, creative, and practical. Analytical intelligence involves abilities tested on traditional IQ tests, such as analyzing, evaluating, and critiquing. Creative intelligence comprises the abilities needed to cope with novel situations, such as creating, discovering, and inventing. Practical intelligence involves abilities needed to solve everyday problems, such as using and applying information. Sternberg has argued that a common set of processes underlies all of these aspects of intelligent behavior: metacomponents, performance components, and knowledge acquisition components. Metacomponents function as a strategy-construction mechanism, arranging the other two types of components into goal-oriented procedures. Knowledge acquisition components are used to obtain new information when no solution to a problem is immediately possible. Performance components do the work of solving the problem. The theory has been applied to diverse cognitive skills and to many different populations.

Production system theories are intended to explain how changes in problem-solving occur. Klahr’s theory explains particularly clearly how self-modifying production systems can advance understanding of development. It focuses on the developing system’s capacity for generalization. In this analysis, generalization includes three components: the time line, regularity detection, and redundancy elimination. The time line is a record of all the situations the system has encountered, its responses to the situations, and the outcomes. Regularity detection operates on the data in the time line to detect repeated patterns. Redundancy elimination looks for parts of procedures that could be eliminated without changing the outcome of processing. Together, these mechanisms allow children to generalize their knowledge to new situations.

Connectionist theories are a class of computer simulation models based on an analogy to the workings of the brain. In them, numerous simple processing units, analogous to neurons, are connected to one another with varying strengths. When presented input, the processing units receive activation from one another, with the processing activity leading to a response. The response is compared to the correct answer, and the strengths of connections are adjusted in ways that would have led to more accurate responding. MacWhinney demonstrated how such a model could learn the German language’s complex system for determining which article should be attached to a given noun, and Shultz demonstrated