It is appropriate to begin a review of research on cognitive development with the work of pioneering researchers such as Luria, Piaget, and Vygotsky, who provided much of the conceptual foundation on which later contributions were built. We will begin with a survey of this legacy, then proceed to more contemporary theories, and finally consider a number of key empirical research topics.

Early Influences

The single most powerful influence on past research into the development of thinking has been the work of Piaget and his collaborators (Inhelder & Piaget, 1958, 1964; Piaget, 1950, 1952, 1953, 1957, 1970), but the influence of Vygotsky (1962) appears to be increasing with time. The work of Luria (1976) deservedly had a major influence on early cognitive development research, but not primarily devoted to thinking. In this chapter, I consider Piaget first, followed by Vygotsky, and then the common ground between them.

Two ideas that were central to Piaget's conception of thought were structure and self-regulation, both of which were also held by the Gestalt school. However, a distinguishing feature of Piaget's theory was that it was based on logico-mathematical concepts, including function, operation, group, and lattice. Although he did not claim that logic defined the laws of thought (cf. Boole, 1854/1951), he used modified logics or “psycho-logics” to model thought.

Piaget's very extensive empirical investigations into the development of infants' and children's cognitions were conceptualized by a succession of distinct logics, which have come to be known as “stages” of cognitive development. The first was the sensorimotor stage, lasting from birth to about one-and-a-half to two years, characterized by structured, organized activity but not thought. During this stage, a structure of actions became elaborated into a mathematical group, meaning that an integrated, self-regulating system of actions developed. Piaget believed that the concept of objects as real and permanent emerged as this structure was
elaborated. The preoperational stage lasted from approximately two to seven years, and during this time semiotic or symbolic functions were developed, including play, drawing, imagery, and language. Thought at this stage was conceptualized in terms of what Piaget called "function logic," the essential idea of which is a representation of a link between two variables. At the concrete operational stage, lasting from eight to about fourteen years, thought was conceptualized in terms of what Piaget called "groupings," which were equivalent to the mathematical concept of a groupoid, meaning a set with a single binary operation (Sheppard, 1978). The essential idea here is the ability to compose classes, sets, relations, or functions, into integrated systems (Halford, 1982). Concepts such as conservation (invariance of quantity, number, weight, and volume), seriation or ordering of objects, transitive inference, classification, and spatial perspectives emerge as a result of the more elaborate thought structures that develop during this time. At the formal operational stage, beginning in adolescence, the ability to compose concrete operations into higher-level structures emerges with the result that thought has greater autonomy and flexibility.

Cognitive development depended, according to Piaget, on assimilation of experience to cognitive structures with accommodation of the structure to the new information. The combination of assimilation and accommodation amounts to a process of self-regulation that Piaget termed "equilibration." He rejected the associationist learning theories of the time, although his conceptions in many ways anticipated modern conceptions of information processing and dynamic systems.

The work of the Piagetian school has been one of the most controversial topics in the field, and claims that Piaget was wrong in many important respects are not uncommon (Bjorklund, 1997; Gopnik, 1996). The following points are intended to help provide a balanced account of this issue. First, Piaget's empirical findings have been widely replicated (Modgil, 1974; Sigel & Hooper, 1968). That is, children have been found to perform as Piaget reported on the tests he used. The major challenges to his findings have been based on different methods of assessment, the claim being that his methods underestimated the cognitive capabilities of young children (Baillargeon, 1995; Bryant, 1972; Bryant & Trabasso, 1971; Donaldson, 1971; Gelman, 1972). However, these claims also have been subject to controversy. Miller (1976) showed that nonverbal assessments did not demonstrate improved reasoning if the cognitive skills employed were taken into account, and a similar point was made about subsequent research by Halford (1989). However, there were also some hundreds of training studies, reviewed by Field (1987) and Halford (1982), that were sometimes interpreted as showing that cognitive development could be accelerated and depended more on experience than on development of thought structures. The stage concept has also been heavily criticized for theoretical inadequacies (Brainerd, 1978) and for lack of empirical support (Bruner, Oliver, & Greenfield, 1966). In particular, acquisition tends to be gradual and experience-based rather than sudden or "stage-like," and the concurrence between acquisitions at the same stage often has not been as close as Piagetian theory might be taken to imply. However, there have also been some spirited defenses of Piaget (Beilin, 1992; Lourenco & Machado, 1996), and Smith (2002) has given a contemporary account of Piagetian theory. See also the special issue of *Cognitive Development* edited by Bryant (2002) on "Constructivism Today."

The underlying problem here seems to have been that it is difficult to operationalize Piagetian concepts in the methodologies that evolved in Anglo-Saxon psychology to about 1970. His conceptions have been more compatible with methodologies that developed after the "cognitive revolution," including information processing and dynamic systems theories. In the next section I consider alternative ways of conceptualizing the development of children's thought.

The work of Vygotsky (1962) was the other major influence on research into the
development of thinking, and his contribution is becoming increasingly influential even today (Lloyd & Fernyhough, 1999). Three of Vygotsky's most important contributions were his ideas on the relation between thought and language, his emphasis on the role of culture in the development of thinking, and the zone of proximal development. Early in the history of cognitive development research, there was considerable debate as to whether thought depends on language development, as implied by Bruner, Olver, & Greenfield (1966), or the reverse, as implied by Slobin (1972). Vygotsky (1962) proposed that thought and language have different origins both in evolution and in development. Language was initially social in character, whereas problem solving was initially motor. Language and thought develop independently for some time after infancy; then the young child develops egocentric speech, the beginning of the representational function. Finally, children develop "inner speech," which serves the symbolic function of thought. Vygotsky emphasized the interaction between biological maturation and social experience. As the child matures, language becomes an increasingly important influence on the development of thought and is the chief means by which culture is absorbed by the child. Vygotsky's concept of the zone of proximal development, which means that new developments are close to existing cognitive abilities, is broadly consistent with Piaget's notion that new knowledge is assimilated to existing structure. This is part of a larger picture in which both Piaget and Vygotsky saw cognitive development as an active organizing process that tends toward an equilibrium with its own internal processes and with the external environment. Piaget's work had greater early influence, but the impact of Vygotsky's work is increasing at what appears to be an accelerating rate. Among the many areas in which it has been important are the development of education theory (Gallimore & Tharp, 1999) and research on collaborative problem solving (Garton, 2004; see also Greenfield, Chap. 27.)

Development of Theory

Theory of development of reasoning diversified in numerous directions in the latter half of the twentieth century and our conceptions of reasoning processes have undergone some fundamental changes. Perhaps one of the most important is that there is much less reliance on logic as a norm of reasoning and more emphasis on the interaction between reasoning processes and the child's experience. Information processing theories were one of the first lines of development following the impact of Piaget and Vygotsky, so it is appropriate to consider them first.

Information Processing Theories

An attempt to conceptualize development of thinking in terms of information processing concepts was made by what became known as the Neo-Piagetian school (Case, 1985, 1992a; Case et al., 1996; Chapman, 1987, 1990; Fischer, 1980; Halford, 1982, 1993; McLaughlin, 1963; Pascual-Leone, 1970; Pascual-Leone & Smith, 1969). These models, reviewed in detail by Halford (2002), reconceptualize Piaget's stages in terms of the information processing demands they make. All of them postulate that higher information processing capacity becomes available with development either through maturation (Halford, 1993) or increased processing efficiency that leaves more capacity available for working memory (Case, 1985). Note that these processes are not mutually exclusive. Chapman and Lindenberger (1989, p. 238) attempted to synthesize these theories under the principle that "the total capacity requirement of a given form of reasoning is equal to the number of operatory variables that are assigned values simultaneously in employing that form of reasoning in a particular task."

Other theoretical developments were more independent of the Piagetian tradition. An important class of theories was based on computer simulations first using symbolic architectures (Halford, Wilson, & McDonald, 1995; Klahr & Wallace, 1976;
Simon & Klahr, 1995) and later using neural nets (Elman, 1990; McClelland, 1995; Shultz, 1991; Shultz, et al, 1995). The model of Klahr and Wallace (1976) was concerned with quantification operators, including subitizing (direct estimation of small sets without counting), counting, and estimation (approximate quantification of large sets such as crowds). It was used to model conservation or understanding that a quantity remains invariant despite transformations of physical dimensions. In a typical simple number conservation task, two rows of beads are placed in one-to-one correspondence. Then one row is transformed (e.g., by spacing objects more widely and thus increasing the length of the row without adding any items); then the child is asked whether each row still contains the same number or whether they are different. Pre-conserving children cannot answer this question correctly because they have not learned that the transformation leaves number invariant. In the model of Klahr and Wallace (1976) the task is performed initially by quantifying first one row followed by the other in the pretransformed display and then comparing the results. The transformed row is quantified again after the transformation and found to be still the same as the other row. With repeated quantification before and after a transformation, the rule that pre- and post-transformed quantities are equal is learned, and the quantification operators are no longer employed. (See also Chap. 17 by Lovett & Anderson, on production system models of thinking.)

The Q-SOAR model of Simon and Klahr (1995) applied Newell's (1990) SOAR architecture to Gelman's (1982) study of number conservation acquisition. Children are shown two equal rows of objects, asked to count each row in turn and say how many each contains, then to say whether they are the same or different. Then one row is transformed and the preconserving child is unable to say whether they are the same or different. This is represented in Q-SOAR as an impasse. The model then searches for a solution to the problem using the quantification procedure of Klahr and Wallace (1976). With repeated experience, the model gradually learns to classify the action of spacing out the items as a conserving transformation, using the learning mechanism of the SOAR model, called "chunking," which has been shown to have considerable generality.

Acquisition of transitive inference was simulated by the self-modifying production system model of Halford, Smith, et al. (1995). Development of transitive inference strategies is guided by a concept of order based on any representation of an ordered set of at least three elements. When no production rule exists for a given problem, the model uses analogical mapping and means-end analysis to determine the correct answer; then a production rule is created to handle that case. Rules are strengthened or weakened by subsequent experiences with success or failure.

**Neural Net Models**

Neural net models of thinking are reviewed by Doumas and Hummel (Chap. 4), but the contribution of neural net models to cognitive development is considered here. A good way to illustrate neural net models of cognitive development is to examine McClelland's (1995) model of children's understanding of the balance scale. The net is shown schematically in Figure 22.1 together with a balance scale problem. It is a three-layered net, which means that activation is propagated from the input units to the hidden (middle) layer and then to the output layer. There are four sets of five input units representing one-to-five weights on pegs one to five steps from the fulcrum on both left and right sides. The units that are activated are shown as black. The activations in the input units represent the problem in the top of the figure. In the first set of input units, representing number of weights on the left, unit 3 is activated, coding the three weights on the left. Similarly, in the second set of input units, representing weights on the right, unit 4 is activating, coding four weights on the right. Distances are coded in a similar way by the two sets of input
units on the right. In the first set, unit 3 is activated, coding the weights on peg 3 on the left, whereas in the second set, unit 2 is activated, coding weights on peg 2 on the right.

There are four hidden units (shown in the middle of the net), two of which compare weights and two that compare distances. The units that are more highly activated are shown as black, although activations would be graded, rather than all-or-none. Finally, there are the output units that compute the balance state. Activation of an output unit represents the corresponding side of the balance beam going down. If the beam is balanced, the activations in the output units would be equal, which is defined as being within 0.3 of each other.

The operation of the unit can be understood from the connection weights between units, which are shown schematically in Figure 22.1 as +/−. The second hidden unit has positive connections to all input units representing weight on the right and negative connections to all input units representing weight on the left (although only a single arrow is shown in each case for simplicity). This unit is more strongly activated because weight on the right is greater than on the left. The first hidden unit has the opposite pattern of weights and will be more strongly activated if weight is greater on the left. The second hidden unit also has positive connections to the right output unit. Thus, greater weight on the right will tend to produce greater activation on the right output unit, representing a tendency for the right side going down. The second pair of hidden units compare distances in corresponding fashion. The activations of the output units depend on activations of hidden units comparing both weights and distances. In this case the greater weight on the right tends to make the right side go down, but this is countered by the greater distance on the left; thus, the predicted position of the beam will be approximately balanced, although, in fact, the left side would go down. The network does not compute the product of weight and distance but compares the influences of weights and distances on each side.

The network was trained by backpropagation; that is, comparing the network’s output on each trial with the correct output and then adjusting the connection weights to reduce the discrepancy. The training would
result in the units representing larger weights or larger distances having greater connection weights to the hidden units. Thus, metrics for weight and distance emerge as a result of training and are not predefined in the net. This is possibly the most important property of the model because it shows how a structured representation can emerge from the process of learning to compute input-output functions that match those in the environment.

The model also captures a number of crucial developmental results. Its progress through training corresponded with the course of development as defined by Siegler's (1981) rules. According to Rule I, judgments are based on weight, irrespective of distance. In Rule II, distance is considered if the weights are first found to be equal. Rule III asserts that weight and distance are considered but difficulty is encountered when weight is greater on one side and distance is greater on the other. Rule IV (torque rule) involves comparing the product of weight and distance on the left side with the product of weight and distance on the right side. The model also captured the torque difference effect – that is, the difference between the product of weight and distance on the left side and on the right side affects children's performance, because they are more likely to recognize that one side will go down if torque difference is large even though there is no logical basis for this given that even a small torque difference will cause one side to go down. This is one of many ways in which neural net models capture psychological properties of task performance.

This model computes the balance state as a function of weight and distance on left and right sides of the balance beam. However, understanding the balance beam also entails determining weight or distance values that will make the beam balance. There are effectively five variables here: \( W_l, D_l, W_r, D_r \), and balance state. Complete understanding of the balance scale would include being able to determine any variable given the other four; that is, compute all five functions implicated by the balance scale concept (Surber & Gzesh, 1984). Other restrictions are that, as Marcus (1998a, 1998b) has pointed out, if the model is trained on two or three weights on either side, it cannot generalize to problems with four or five weights. Again, however, it would be reasonable to expect that children would generalize in this way. The conclusions therefore are that the model can be trained to compute one function implicated by the balance scale, albeit under restricted conditions, and that it does not fully capture understanding of the concept but is nevertheless an important step forward in our understanding of cognitive development because it shows how structured representation can emerge.

The balance scale model by McClelland (1995) is a three-layered, or backpropagation, net. This type of architecture has been used in a great many models, in cognitive development and elsewhere. One reason is that it can, in principle, compute any input-output function. The simple recurrent net (Elman, 1990) is an important model in this class. In this type of net, activations in the hidden units are copied over into context units. On the next trial, activations in the hidden units are influenced by activations in both the input units and the context units. The result is that the output of the net is influenced by representations on previous trials as well as by the current input. The net therefore takes account of links between events in a sequence. The model was trained to predict the next word in a sentence. Training was based on a large corpus of sentences by representing each successive word in the input units, and the output units were trained to represent the next word. Feedback was given concerning the accuracy of the output, thereby adjusting the connection weights to improve the model's prediction. The model learned to predict the next word in a sentence and respected grammatical categories even when words in related categories spanned embedded clauses. Cluster analysis of the hidden unit activations showed that words in the same grammatical category, such as nouns or verbs, tended to have similar activations. Semantically similar words, such as those
representing animals or foods, also tended to have similar hidden unit representations. Elman (1990) was careful not to predefine categories, and the inputs used were orthogonal; thus, no pre-existing similarities were supplied to the model. Similarities were created in the hidden unit activations that reflected the input-output functions the model was required to learn. Therefore, to the extent that categories developed, they are an emergent property of the model and one that reflects contingencies in the environment. This model, like that of McClelland (1995), offers a possible mechanism by which structured representation might be acquired.

The ability of simple recurrent nets to predict sequences has been utilized to model infants' expectations of the reappearance of occluded objects (Mareschal, Plunkett, & Harris, 1995; Munakata et al., 1997) thereby simulating infants' understanding of the object concept (Baillargeon et al., 1990). These models are basically consistent with the model of Smith et al. (1999). Again, however, there have been limitations. Marcus (1998a, 1998b) found that the model of Munakata et al. (1997) did not generalize to objects in new positions on the display.

The potential of models such as this to learn regularities in the environment and acquire concepts has inspired a whole new approach to cognitive development (Elman et al., 1996). Elman et al. (1996) see connectionism as giving more powerful means to analyze the gene-environment interactions that are the basis of development. They advocate a form of connectionism that is founded in biology, is influenced by developmental neuroscience, and that can produce neurologically plausible computational models. Although they see an undoubted role for inateness in cognitive development, they argue some nativist conceptions underestimate the potential for new cognitive forms to emerge from the interaction of neural processes. The simple recurrent net nicely illustrates how representations that respect distinctions between word categories emerge from the model's interaction with the environment.

Cascade Correlation Models

Cascade correlation models provide a mechanism by which the dimensionality of representations can be increased to handle increased dimensions in the task. They do this by adding units to the hidden layer. The initial net has minimal hidden units and sometimes starts with none. Training takes place in two modes. In the first mode, weights are adjusted to yield the appropriate output for each input. In the second mode, hidden units are recruited to increase the accuracy of the output. Recruitment is based on correlation between a candidate's activation and the existing error of the network. After recruitment of a hidden unit, training continues in the first mode and the system cycles between the modes until a learning criterion is reached.

Cascade correlation models have been used to model a number of developmental phenomena (Shultz, 1991; Shultz et al., 1995; Sirois & Shultz, 1998). Shultz and colleagues used cascade correlation to model the same balance scale problem modeled by McClelland (1995). The initial net was similar to that used by McClelland and shown in Figure 22.1, but without hidden units. Initial training was with problems varying only in weight, and the net performed consistent with Siegler's (1981) Rule 1. Once the distance variable was introduced, the net recruited a single hidden unit. It then progressed to Rule 2 and higher rules that take account of distance, effectively simulating the developmental progression in a manner similar to McClelland's model (1995).

Neural Net Models and Symbolic Processes

Concern that three-layered net models do not capture symbolic processes has been expressed by Fodor and Pylyshyn (1988). Properties that are considered essential by Fodor and Pylyshyn (1988) are compositionality and systematically. The essential idea of compositionality is that symbols must retain their identity and their meaning when combined into more complex representations. Thus, the cognitive symbols for "dog"
and "happy" must retain their identity when combined into the symbol for "happy dog." Prototypes are not necessarily compositional in this way (Fodor, 1995). One problem for three-layered net models is that the representations in the hidden units do not necessarily include the components of the input in a form that is recognizable to the performer. Any structure that exists in the hidden layer must be discovered by an external observer (the experimenter) using techniques such as cluster analysis (Elman, 1990). Representations in hidden units are not accessible to strategic processes. They are more like implicit knowledge (Karmiloff-Smith, 1994).

Systematicity, in essence, means that cognitive processes are subject to structural constraints independent of content. Three-layered nets lack strong systematicity (Marcus, 1998a, 1998b; Phillips, 1994), meaning they cannot generalize to an element that has not occurred in the same role before, even if the element is familiar. Thus, a net trained on "John loves Mary" and "Tom loves Jane" could generalize to "John loves Jane" but not to "Jane loves John" or even to "Mary loves John." Nets of this type learn representations that are needed to compute the input-output functions on which they are trained, but they do not learn abstract relations.

Although three-layered net models have real potential to advance research on cognitive development (Bray et al., 1997), it appears they lack the structural properties that have long been regarded as characteristic of higher cognition (Chomsky, 1980; Humphrey, 1951; Mandler & Mandler, 1964; Miller, Galanter, & Pribram, 1960; Newell, 1990; Piaget, 1950; Wertheimer, 1945). One response to this problem (Smolensky, 1988) is that neural net models seek to explain symbols as emergent properties of more basic processes, as was illustrated earlier. A second approach has been to develop symbolic neural net models of higher cognitive processes (Doumas & Hummel, Chap. 4; Shastri & Ajjanagadde, 1993; Smolensky, 1990; but see also Halford, Wilson, & Phillips, 1998). A symbolic connectionist account of cognitive development has been given by Halford and his collaborators (Halford, 1993; Halford, Wilson, & Phillips, 1998). See Halford (2002) for a summary of this approach.

Dynamic Systems Models

Dynamic systems models (Fischer & Bidell, 1998; Fischer & Pare-Blagoev, 2000; van Geert, 1991, 1998, 2000) have offered new ways to analyze developmental data. A dynamic system is a formal system, the state of which depends on its state at a previous point in time. The dynamic system model of van Geert (1998) was designed around principles derived from the work of Piaget and Vygotsky and has a number of interesting properties. It can account for different types of cognitive growth, such as slow linear increase and sudden discontinuities, within the same system. It can also show how a complex, self-regulating system can emerge from the interaction of a few variables. The model was fitted to a number of developmental data sets, and some important developmental phenomena, including conservation acquisition, were simulated. Links have also been made between dynamic systems models and neural net models.

Dynamic systems models have also been linked to other issues. Raijmakers, van Koten, and Molenar (1996) analyzed McClelland's (1993) neural net model of the balance scale and found no evidence of the flags indicating discontinuities that are found in empirical data. They suggest that backpropagation models simulate the type of stimulus-response associations that are characteristic of animals and young children but do not simulate the rule-governed behavior characteristic of older children and adults. In many respects, this finding is consistent with the analysis of the model presented earlier. On the other hand, backpropagation models incorporate learning functions that have been missing from models of higher cognitive processes. As we have seen, they show how structured representations begin to emerge as a result of learning input-output functions.

Although there are acknowledged difficulties with dynamic systems models (van
Geert, 1998), they provide much more sophisticated implementations of important developmental theories, including that of Piaget and Vygotsky. This does not mean that Piaget and Vygotsky are fully vindicated by dynamic systems models, but concepts such as equilibration and self-regulation, which are at the core of their theories, do seem to have a new lease on life. Most importantly, dynamic systems models have potential to deepen our understanding of cognitive developmental processes. And, as Fischer and Pare-Blagoev (2000) point out, there are tools based on Lotus 123 or Microsoft Excel that make dynamic system modeling more accessible.

**Links to Brain Development**

The finding by Thatcher, Walker, and Giudice (1987) of brain growth spurts that appeared to correspond to stage transitions in cognitive development stimulated considerable interest in the explanatory potential of neural maturation. One of the important landmarks in infant development is the A not-B error: If infants are shown a toy hidden at A several times and allowed to retrieve it and then see it hidden at B, before approximately 12 months of age they tend to search for it at A. Studies by Diamond (1988) and Goldman-Rakic (1987) showing the link between frontal lobe function and the A not-B error were important stimuli to work on infant brain development. Case (1992a, 1992b) and Fischer (1987; Fischer & Rose, 1996) have drawn interesting parallels between cognitive development and the growth of connections between the frontal lobes and other brain regions. Robin and Holyoak (1995) and Waltz et al. (1999) have also drawn attention to the role of the frontal cortex in processing relations of the kind described by Halford and his collaborators (Halford, 1993; Halford, Bain et al., 1998; Halford, Wilson, & Phillips, 1998). In a different context, Rudy, Keith, and Georgen (1993) present evidence that configural learning (e.g., conditional discrimination, in which a cue-response link is reversed on change of background) depends on maturation of the hippocampus.

At a more general level, Quartz and Sejnowski (1997) have argued that synaptic growth, axonal arborization, and dendritic development play a role in processing capacity increase with age. They also point out that neural plasticity would cause capacity to increase as a function of experience. This implies that the issue of whether cognitive development depends on capacity, knowledge, or both may need to be redefined. It might be that cognitive development depends on growth of capacity, which is at least partly produced by experience.

**Strategy Development**

Problem-solving strategies are important to reasoning in children and adults, and much of the improvement in children’s reasoning can be attributed to development of more powerful strategies. It is appropriate therefore that much research has been devoted to development of strategies. Following work on rule assessment (Briars & Siegler, 1984; Siegler, 1981), Siegler and his collaborators conducted an extensive study of strategy (Siegler, 1999; Siegler & Chen, 1998; Siegler & Jenkins, 1980; Siegler & Shipley, 1995; Siegler & Shrager, 1984). Two of the models were concerned with development of addition strategies in young children. When asked to add two single-digit numbers, they chose between a set of strategies including retrieving the answer from memory, decomposing the numbers (e.g., $3 + 5 = 4 + 4 = 8$), counting both sets (counting right through a set of three and a set of five, perhaps using fingers), and the *min strategy* of counting on from the top number in the larger set (e.g., $5, 6, 7, 8$, so $3 + 5 = 8$).

Siegler and Shrager’s early strategy choice model (1984) was based on distribution of associations. The idea is that each addition sum is associated with answers of varying strengths, and so for a given sample of children, $2 + 1$ might yield the answer “$3$” 80% of the time; “$1$” or “$2$,” 4%; “$4$,” 3%; and so on. The chance of an answer being chosen is a function of its associative strength relative
to competing answers. The more peaked the distribution, the more likely it will be that a single answer will occur. However, it will be adopted only if it is above the confidence criterion. If not, alternative strategies, such as counting, are sought.

In their later work, Siegler and his collaborators developed the Adaptive Strategy Choice Model (ASCM, pronounced "Ask-em") which makes more active strategy choices. At the beginning, ASCM knows only the small set of strategies typically used by 4-year-olds, but it has general cognitive skills for choosing and evaluating strategies. The model is trained on a set of elementary addition facts; then the min strategy is added to the model's repertoire. This entails counting on from the larger number to be added, so if the sum is $5 + 3$, the procedure is to count $5, 6, 7, 8$. The model chooses a strategy for each problem on the basis of the past speed and accuracy of the strategy and on similarity between the current problem and past problems in which a strategy has been used. Each time a strategy is used, the record of its success is updated, and the projected strength of the strategy for that problem is calculated. The strength of association between a problem and a specific answer is increased or decreased as determined by the success of the answer. One of the strengths of the model is that it can account for variability both between children and between different strategies used by the same child for a particular class of problems. Most importantly, it provides a reasonably accurate account of strategy development in children as they age.

**Complexity**

Children become capable of more complex reasoning with age, and it is therefore important to have some way of comparing the complexities of reasoning tasks. A conceptual complexity theory and accompanying metric that discriminate tasks of different difficulty and explain why they differ is essential to understanding cognitive development. It is also necessary to define *equivalence* in cognitive tasks. In the past, tasks have tended to be regarded as equivalent if they require the same knowledge domain and are similar methodologically, or if they have similar difficulties on a psychometric scale. Although these criteria have great utility, they have not led to an understanding of factors that underlie complexity, nor do they explain why tasks that differ in content or procedure can be of equivalent complexity whereas tasks that are superficially similar can be very different in complexity. Without a means of assigning cognitive tasks to equivalence classes with common properties and relating tasks in different classes to each other in an orderly way, psychology is in a position similar to that of chemistry without the periodic table (Frye & Zelazo, 1998). Two metrics for cognitive complexity have been developed in the past decade.

**Cognitive complexity and control (CCc) theory**

Frye, Zelazo, & Palfai (1995; Zelazo & Frye, 1998) analyze complexity according to the number of hierarchical levels of rules required for the task. A simple task entails rules that link an antecedent to a consequent, $a \rightarrow c$, whereas complex tasks have rules that are embedded in a higher-order rule that modifies the lower level rules; thus, another level is added to the hierarchy. The dimensional change card sort task has been a fruitful implementation of this theory. In a simple sorting task, a green circle might be assigned to the green category and a red triangle to the red category, where categories are indicated by templates comprising a green triangle and a red circle. In a complex task, sorting depends on whether the higher order rule specifies sorting by color, as just mentioned, or by shape. If sorting is by shape, the green circle is sorted with the red circle, and the red triangle is sorted with the green triangle. Normative data (e.g., Zelazo & Frye, 1998; Zelazo & Jacques, 1996) indicate that children typically process a single rule by two years of age, a pair of rules by three years, and a pair of rules embedded under a higher order rule by four years. The dimensional change card sort task has been a useful predictor of other cognitive
performances such as concept of mind (Frye Zelazo, & Palfai, 1995).

**THE RELATIONAL COMPLEXITY (RC) METRIC**

Halford (Halford, Wilson, & Phillips, 1998) defines complexity as a function of the number of variables that can be related in a single cognitive representation. This corresponds to the arity, or number of arguments (slots) of a relation (an n-ary relation is a set of points in n-dimensional space). Normative data indicate that quaternary relations (four related variables) are the most complex that can be processed in parallel by most humans, although a minority can process quinary relations under optimal conditions. Children can process unary relations at one year, binary relations at two years, ternary relations at five years, and the adult level is reached at 11 years (median ages).

Complex tasks are segmented into components that do not overload capacity to process information in parallel. However, relations between variables in different segments become inaccessible (just as a three-way interaction would be inaccessible if two-way analyses were performed). Processing loads can also be reduced by conceptual chunking, which is equivalent to compressing variables (analogous to collapsing factors in a multivariate experimental design). For example, velocity = distance/time but can be recoded to a binding between a variable and a constant (e.g., speed = 80 kph) (Halford, Wilson, & Phillips, 1998. Section 3.4.1). Conceptual chunking reduces processing load, but chunked relations become inaccessible (e.g., if we think of velocity as a single variable, we cannot determine what happens to velocity if we travel the same distance in half the time). Complexity analyses are based on the principle that variables can be chunked or segmented only if relations between them do not need to be processed. Tasks that impose high loads are those in which chunking and segmentation are constrained.

CCC and RC theories have some common ground, but whereas CCC attributes complexity to the number of levels of a hierarchy, RC attributes it to the number of variables bound in a representation. Therefore RC theory is directly applicable both to hierarchical and nonhierarchical tasks. Also, the principles of segmentation and conceptual chunking imply that difficult tasks are those that cannot be decomposed into simpler tasks. In the sorting task discussed with respect to CCC theory, it is necessary to keep in mind that we are sorting by color in order to determine that the green circle is sorted with the green triangle. This means the task cannot be decomposed into two subtasks that are performed independently, because the conflicting dimension is always present.

Andrews and Halford (2002) showed that with four- to eight-year-old children, in the domains of transitivity, hierarchical classification, cardinality, comprehension of relative clause sentences, hypothesis testing, and class inclusion, a single relational complexity factor accounted for approximately 50% of variance and factor scores correlated with fluid intelligence ($r = .79$) and working memory ($r = .66$).

**Increased Dimensionality**

Taking account of extra dimensions is a fundamental requirement for cognitive development. For example, the progression from an undifferentiated concept of heat to a concept that distinguishes heat and temperature entails taking account of the dimensions of mass and specific heat: Heat = temperature x specific heat x mass. Similarly, the distinction between weight and density depends on taking into account volume and specific gravity: Weight = specific gravity x volume. Taking account of the extra dimensions enables children to progress from undifferentiated concepts of heat or weight to more sophisticated concepts that recognize the distinction between heat and temperature or between weight and specific gravity. Thus, they become capable of recognizing that a piece of aluminium weighs less than a similar volume of lead, but a sufficiently large piece of aluminium can weight more than a piece of lead. Arguably, the progression that children make here parallels the development of these concepts in the history of
science (Carey & Gelman, 1991). In an entirely different context, acquisition of conservation of continuous quantity arguably entails taking account of height and width of containers, rather than fixating on height alone (Piaget, 1950). The essential point here is that cognitive representations must include sufficient dimensions to take account of the variations in a phenomenon, so children must represent volume and specific gravity to take account of variations in weight, and so on.

The importance of cascade correlation models, considered earlier, is that they offer a possible mechanism by which extra dimensions can be added to cognitive representations to take account of variations in the task. The model does not have to be told what dimensions to include. It creates dimensions in its own representations, contained in the hidden units, as required for input–output functions on which it is trained. This can be seen as modeling the increased dimensionality of children's cognitive representations as they learn to predict variations in the environment. This mechanism illustrates the potential for neural nets to provide the long-hoped-for basis of constructivism without postulating that all the dimensions children attend to are innately determined (Elman et al., 1996, but see also Marcus, 1998a, 1998b). Mareschal & Shultz (1996) suggest that cascade correlation models can provide a way to increase the computational power of a system, thereby overcoming a criticism by Fodor (1980) of constructivist models of cognitive development.

Knowledge and Expertise

The theories considered so far have placed major emphasis on development of reasoning processes, but acquisition and organization of knowledge is equally important. Furthermore, knowledge acquisition interacts with development of reasoning processes to determine how effectively children can reason and solve problems.

Several important lines of research have recognized acquisition of knowledge as a major factor in cognitive development (Carey & Gelman, 1991; Ceci & Howe, 1978; Keil, 1991). Cognitive development can also be seen as analogous to acquisition of expertise, so the reasoning of young children is analogous to that of the novice in a domain. The effect of domain expertise on even the most basic cognitive functions was demonstrated by Chi (1976), who showed that child chess experts outperformed adult chess novices on a simple recall test of chess pieces on a board. On recall of digits, the children performed according to age norms, and well below the level of adults. This experiment cannot be interpreted validly as showing that memory capacity does not change with age, because capacity is not measured and the experiment is quite consistent with an increase in capacity with age that is overridden by differences in domain expertise. The capacity question requires quite a different methodology. However, the study does show how powerful effects of domain knowledge can be. Carey (1991) argues that differentiation of heat and mass by young children is similarly attributable to knowledge acquisition. There is, of course, no logical reason to assume that explanations based on knowledge acquisition are necessarily incompatible with explanations based on growth in capacity. Most of the evidence suggests an interaction of these processes.

Although not in the mainstream of knowledge research in cognitive development, Halford and Wilson (1980) and Halford, Bain et al. (1998) investigated possible mechanisms for acquisition of structured knowledge along lines similar to the induction theory proposed by Holland et al. (1986). See also a special issue of Human Development (Kuhn, 1995) on reconceptualizing the intersection between development and learning.

Advances in our understanding of children's knowledge have had a pervasive influence on research in the field, and it would be hard to think of a domain that has not been touched by it. In this review, knowledge is considered in relation to children's expertise in specific domains, including conservation, transitivity, classification, prototype formation, theory of mind, and scientific and
mathematical concepts. (See also Chap. 14 by Novick & Bassok on problem solving and expertise.)

**Domain Specificity versus Generality**

The view that cognitive processes are domain-specific rather than domain-general has developed in parallel with knowledge acquisition theories of cognitive development and has been reinforced by Fodor's proposal (1983) that many cognitive processes are performed by specialized modules. For example, it has been proposed that conditional reasoning (i.e., reasoning in which the major premise has the form “if-then”) might depend on a module for cheater detection (Cheng & Holyoak, 1989; Cosmides & Tooby, 1992; but see Cosmides & Tooby, 1989), that understanding mathematics might depend on innate enumeration processes (Gelman, 1991), or that reasoning about cause might be facilitated by a module for processing causal information (Leslie & Keeble, 1987). One achievement has been to show that young children understand the distinction between artifacts and natural kinds (Keil, 1991) and have considerable knowledge of basic facts about the world. For example, they understand that animals move autonomously, have blood, and can die (Gelman, 1990; Keil, 1995). The distinction between animate and inanimate objects even seems to be appreciated in infancy (Gergely et al., 1995). One result of these developments has been an increasing biological perspective in theories of children’s reasoning (Kenrick, 2001). Domain-specific knowledge must now be seen as having a major influence on the developing cognitions of children, but it does not displace domain-general knowledge entirely. Basic cognitive operations such as memory retrieval, and basic reasoning mechanisms such as analogy and means-end analysis, are applicable across domains. Furthermore, some higher reasoning processes such as transitive inference and classification are found to correspond across domains (Andrews & Halford, 2002). Theories such as that of Case (1985; Case et al., 1996) recognize the importance of both domain-specific and domain-general processes.

**Reasoning Processes**

Piaget based his theory of cognitive development on the child’s progression through increasingly complex logics, but this approach has not been generally successful as a way of modeling children’s reasoning (Halford, 1993; Gershon, 1974). Considerable success has been achieved in accounting for adult reasoning using mental models (Johnson-Laird & Byrne, 1991), analogies (Gentner & Markman, 1997; Hofstadter, 2001; Holyoak & Hummel, 2001; Holyoak & Thagard, 1995), schemas (Cheng et al., 1986), and heuristics (Kahneman, Slovic, & Tversky, 1982).

Analogical reasoning is reviewed by Holyoak (Chap. 6), but the implications for understanding cognitive development are considered here. An analogy is a structure-preserving map from a base or source to a target (Gentner, 1983; Holyoak & Thagard, 1989). The map is validated by structural correspondence rather than similar elements. Structural correspondence is defined by two principles; **uniqueness of mapping** implies that an element in the base is mapped to one and only one element in the target; **symbol-argument consistency** implies that if a relation symbol $r$ in one structure is mapped to the relation symbol $r'$ in the other structure, the arguments of $r$ are mapped to the arguments of $r'$ and vice versa. These principles operate as soft constraints and can be violated in small parts of the mapping if the overall mapping conforms to the criteria. Success in mapping depends on representation of the corresponding relations in the two structures and on ability to retrieve the relevant representations, which, in turn, depends on knowledge of the domain. Research on children’s analogical reasoning is reviewed by Goswami (1998, 2002).

Numerous studies have assessed young children’s ability to perform simple proportional analogy — that is, problems of the form A is to B as C is to D. Brown (1989)
showed that children as young as three years could use analogies for both learning and problem solving if they understood the relevant relations, were able to retrieve them from memory, and understood the aims of the task. This was borne out by Goswami (1989), who showed that three-, four-, and six-year old children could perform analogies based on relations they understood, such as cutting or melting (e.g., chocolate:melted chocolate::snowman:melted snowman). In a less tightly structured context, Gentner (1977) showed that four- to six-year-old children could map human body parts to inanimate objects such as trees (e.g., if a tree had a knee it would be on the trunk a short distance above the ground). There appears to be consensus now that young children can perform analogies with simple relations if they have the relevant domain knowledge and if the test format is appropriate to the age of the children.

Young children can also use analogies for problem solving (Brown, Kane, & Echols, 1986; Crisafi & Brown, 1986; Holyoak, Junn, & Billman, 1984). In the study by Holyoak et al. (1984), children were told a story about a genie who transferred jewels from one bottle to another by rolling his magic carpet into a tube and rolling the jewels down it. Then they were given the problem of transferring gumballs from one jar to another using a tube made by rolling a sheet of heavy paper. Even four-year-olds showed evidence of analogical reasoning. Gholson et al. (1996) tested children from first to fifth grade on transfer from missionaries and cannibals problems to jealous husbands problems, both of which require a sequence of moves to be selected for transferring people from one place to another without violating constraints. In a second experiment they used similar problems that required a sequence of arithmetic steps to be chosen. The children showed evidence of analogical transfer, based on representation of common relations. Pauen and Wilkening (1997) found evidence that second- and fourth-grade children transferred selected aspects of balance scale problems to simple physical force problems.

Mental models have been found effective for providing explanations of human reasoning (Johnson-Laird & Byrne, 1991; Johnson-Laird, Chap. 9). A mental model is more content-specific than a logical rule and is used by analogy. Gentner and Gentner (1983) showed that high school and college students could use water flowing in pipes as mental models of electricity, so pipes were mapped to conductors, constrictions in pipes were mapped to resistors, water pressure to voltage, water flow to electric current, and reservoirs to batteries. Furthermore, reservoirs placed above one another were mapped into batteries in series, and the increase in water pressure was mapped to the increase in voltage, and so on.

It appears that mental models are also an effective way of accounting for development of reasoning in children (Barrouillet, Grosset, & Lecas, 2000; Barrouillet & Lecas, 1999; Halford, 1993; Markovits & Barrouillet, 2002). Marcovits and Barrouillet (2002) have developed a mental models theory that accounts for most of the data on children's conditional (if-then) reasoning. Conditionals may refer to classes (e.g., if X is a dog then X is an animal, or simply, all dogs are animals) or to causal relations (e.g., if it rains, the ground will get wet). For the problem, if p then q, p therefore q (modus ponens), construction of a mental model begins with the following representation:

\[
p \quad q
\]

This represents the case in which p and q are both true. The model could now be fleshed out with other possibilities as follows (where \( \neg \)p is read as “not p”):

\[
p \\
\neg p \\
\neg p \\
\neg p \\
q
\]

The second premise “p” is processed by selecting those components of the model where p is true, in this case, the first line; then inference is made by examining these cases. In this model, in the only case in which
p is true, q is also true; thus, the inference is "q." For example, if the major premise were "if an animal is a dog then it has legs," then the initial model would be

\[
\begin{align*}
dog & \quad \text{legs} \\
\text{not dog} & \quad \text{no legs} \\
\text{not dog} & \quad \text{legs}
\end{align*}
\]

This could be fleshed out with alternative cases such as

\[
\begin{align*}
\text{not dog} & \quad \text{no legs} \\
\text{not dog} & \quad \text{legs}
\end{align*}
\]

Thus the premises are processed as relational propositions, referring to specific instances, and are fleshed out by retrieving relevant information from semantic memory. The accuracy of children's reasoning depends on the fleshing out process, which is influenced by availability of relevant information in memory and by working memory capacity. The minor premise "not dog" can produce the fallacious inference "no legs" (denial of the antecedent) if the second line of the mental model is missing. This could occur if the child failed to retrieve any cases of things that are not dogs but have legs. Similarly, the minor premise "legs" can produce the fallacious inference "dog" (affirmation of the consequent). Markovits (2000) has shown that children are more likely to recognize that these inferences are not justified if they can readily generate the alternative cases. In the aforementioned example, it is easy to generate instances of things that are not dogs but have legs. In a problem such as "if something is a cactus then it has thorns," generation of alternative cases is more difficult, and children are less likely to recognize the fallacies. The second major factor is processing (working memory) capacity. More complex problems entail representation of more relations. The example just given effectively entails three relations corresponding to the three lines of the mental model. A simpler problem would consist of only the first and second lines and corresponds to a biconditional interpretation of the major premise. This representation is simpler. Increase in effective capacity with age enables children to reason correctly on more complex problems. The model of Marcovits and Barrouillet (2002) can handle both content effects (Barrouillet & Lecas, 1998; Leevers & Harris, 1999) and complexity effects (Halford, Wilson, & Phillips, 1998).

Other studies of children's conditional reasoning have utilized the Wason selection task (Wason, 1966; see also Evans, Chap. 8). The task entails four cards containing (say) an A, B, 4, and 7, and participants are told that there is a letter on one side of each card and a number on the other. They are asked which cards must be turned over to test the proposition that if there is an A on one side there must be a 4 on the other. The correct answer, cards containing the A and 7, is rare even among adults. There are well-known content effects, and it has been shown that versions of the task based on permission (Cheng et al., 1986) or cheater detection (Cosmides & Tooby, 1992) are performed better. Similar improvements have been observed in children (Cummins, 1996; Light et al., 1986).

The literature supports the claim that conditional reasoning is possible for children, even as young as four, and improvements can be produced by more appropriate task presentation (Markovits et al., 1996) and by experience, but considerable development occurs throughout childhood (Muller, Overton, & Reene, 2001). Among the relatively late-developing competences are understanding of logical necessity (Falmagne, Mawby, & Pea, 1989; Kuhn, 1977; Morris & Sloutsky, 2002; Osherson & Markman, 1975) and reasoning that requires representation of complex relations.

**Elementary Concepts**

Conservation, transitivity, and classification are three concepts that have long been considered fundamental to children's reasoning. All have been controversial because Piaget's claim that they are concrete operational and unattainable before seven to eight years has been contested by many researchers. We briefly consider each in turn.
Conservation

Perhaps the most widely researched concept in the field, conservation, is still not well understood. The Q-SOAR model of conservation acquisition was briefly reviewed earlier, and other models exist (Caroff, 2002; Halford, 1970; Shultz, 1998; Siegler, 1995). However, the issues that have received most attention in the literature concern how conservation should be measured and the age at which children master it. A number of authors have argued that the Piagetian tests misled children and therefore underestimated their understanding. A common cause of the alleged misunderstanding is that an increase in the length of a row of objects (or an increase in the height of a column of liquid) makes the number (or amount) appear greater. This received substantial support from a study by Gelman (1969), who used an oddity training procedure to induce five-year-old children to attend to number rather than length and showed that they conserved number. This interpretation received further support from McGarrigle and Donaldson (1975), who improved conservation in children aged four to six years by having a "naughty teddy" perform the transformation, thereby making it accidental and removing any suggestion that an increase in amount was intended. These studies, like a host of others, showed improved performance in children aged four to six years by having a "naughty teddy" perform the transformation, thereby making it accidental and removing any suggestion that an increase in amount was intended. These studies, like a host of others, showed improved performance in children about five to six years of age. However, Bryant (1972) eliminated length cues and showed that three- and four-year-old children carried a pretransformation judgment over into the posttransformation situation. However, his claim that this demonstrated conservation was disputed by Halford and Boyle (1985). Sophian (1995) also failed to replicate Bryant's finding of early conservation and showed that conservation was related to understanding of counting, suggesting that conservation reflects some aspects of children's quantitative concepts. From extensive reviews of the conservation literature (Halford, 1982, 1989), it seems there is clear evidence of conservation at approximately five years of age, which is earlier than Piaget claimed, but not as early as claimed by Bryant (1972).

Transitivity and Serial Order

A transitive inference has the form: aRb, bRc implies aRc, if R is a transitive relation. For example, a > b, b > c implies a > c. Piaget's claim of late attainment was challenged by Bryant and Trabasso (1971), who trained three- to six-year-old children to remember the relative lengths of adjacent sticks in a series (e.g., a < b, b < c, c < d, d < e). Then they were tested on all possible pairs. The crucial pair is b > d because this was not learned during training and must be inferred from b < c, c < d. Also, the bd pair avoids the end elements, which tend to be labeled as small (a) or large (e). Bryant and Trabasso found that three- and four-year-old children performed above chance on the bd pair, suggesting that they made a transitive inference. Riley and Trabasso (1974) showed that both children and adults performed the task by ordering the elements—that is, they formed the ordered set a, b, c, d, e. This in itself does not affect the validity of the test because an asymmetric, transitive binary relation is a defining property of an ordered set, so the children presumably utilized transitivity in some way while ordering the elements. The problem, however, was that, to facilitate acquisition, the premise pairs were presented initially in ascending or descending order (a < b, b < c, c < d, d < e, or the reverse). This clearly gave children undue help in ordering the elements. Furthermore, children who failed to learn the premise pairs were eliminated, and elimination rates were as high as 50% in some experiments. The problem with this is that children might have failed to learn the premises because they could not determine the correct order, which, in turn, might reflect lack of understanding of transitivity. When Kallio (1982) and Halford and Kelly (1984) eliminated these extraneous sources of help, success was not observed below five years of age. Subsequent research (Andrews & Halford, 1998; Pears & Bryant, 1990) has confirmed that transitive inference is
understood by only a minority of four-year-olds, and the median age of attainment is about five years. For more extensive reviews and theoretical discussions, see Brainerd & Reyna (1993), Breslow (1981), Thayer & Collyer (1978), and Halford (1982, 1993).

A derivative of the Bryant and Trabasso (1971) paradigm, transitivity of choice, has found wide use in animal studies (Boysen et al., 1993; Chalmers & McGonigle, 1984; von Fersen et al., 1991; McGonigle & Chalmers, 1977; Terrace & McGonigle, 1994). Participants are trained to choose one member of each pair in a series. For example, they are rewarded for choosing A in preference to B, B in preference to C, C in preference to D, and D in preference to E. Transitivity of choice is indicated by choice of B in preference to D. However, whereas transitive inference implies an ordinal scale of premise elements, in transitivity of choice there is no such scale (Markovits & Dumas, 1992). Furthermore, whereas the transitive inference task is performed dynamically in working memory, following a single presentation of premises, the premise pairs in transitivity of choice are learned incrementally over many trials, and the task can be performed by associative processes (Wynne, 1997). Although both paradigms are important, transitive inference and transitivity of choice should not be regarded as equivalent tests of the transitivity concept.

Classification

Concepts and categories are reviewed by Medin and Rips (Chap. 3). Developmentally, categorization appears to progress from prototypes, arguably the most basic form of categorization, to more advanced categories, including those based on rules or theories. All advanced categories appear to have a label or symbol (e.g., "dog" for the dog category) so it will be convenient to deal with them under the heading of symbolic categories.

Prototype Models of Classification

There is evidence that infants can form prototypic categories (Rosch & Mervis, 1975; Rosch et al., 1976) and can recognize that a set of objects with similar features, such as animals (dogs, cats), form categories (Quinn, 2002). They are also sensitive to the correlation between attributes (Younger & Fearing, 1999). However, prototypes are arguably subsymbolic because they are well simulated by three-layered nets (Quinn & Johnson, 1997) and do not have properties such as compositionality that are basic to symbolic processes (Fodor, 1995; Halford, Phillips, & Wilson, unpublished manuscript). Mandler (2000) argues that infants make the transition from perceptual categories, which enable objects to be recognized by their appearance, to conceptual categories, defined by the role objects play in events and that serve as a basis for inductive inference.

Neural nets are a very suitable basis for constructing models of prototype formation, and McClelland and Rumelhart (1985) produced an early prototype model that fundamentally changed the way we view categorization. Quinn and Johnson (1997) developed a three-layered net model of prototype formation in infants. There were thirteen input nodes that encoded attributes of pictorial instances of four animals (cats, dogs, elephants, rabbits) and four kinds of furniture (beds, chairs, dressers, tables). There were three hidden units and ten output units, two of which coded for category (animals, furniture) and the remainder coded for the eight instances. After the net was trained to recognize categories and instances, the representations in the hidden units were examined. Initially there was no differentiation, then the units differentiated mammals and furniture, then instances were distinguished within the categories. The study is important for showing how categories can be formed by a learning algorithm. As with the models of McClelland (1995) and Elman (1990) discussed earlier, the representations emerge from the learning process.

Symbolic Categories

Even young children form categories based on, and draw inductive inferences about, essential or nonobvious properties (Gelman,
This means that objects are categorized on the basis of hidden properties that cause their surface or observable features, so animals contain essential biological material that enable them to move, eat, make characteristic sounds, and reproduce animals of the same kind. Categorization by essential properties is sometimes interpreted as evidence that people have theories about the domain, such as theories about the nature of animals, although there are also interpretations based on causal laws rather than essences (Rehder & Hastie, 2001; Strevens, 2000). There is strong evidence that young children can make inductive inferences based on category membership. Gelman and Markman (1986) presented four-year-olds with a picture of a bird, told the children a property of the bird (feeds its young with mashed up food) and found that the children attributed the property to other, even dissimilar, birds but not to a different category such as bats. Young children appear to generalize even nonobservable properties on the basis of category membership, independent of appearance.

The ease with which young children make inductive inferences about categories contrasts with the difficulty they have in reasoning about hierarchically structured categories. For example, given twelve apples and three oranges, when asked whether there are more apples or more fruit, they tend to say there are more apples. This task is a derivative of class inclusion items originally used by Inhelder & Piaget (1964). The Piagetian hypothesis was that children lacked a concept of inclusion till they reached the concrete operations stage, but many alternative hypotheses have been proposed (McGarrigle, Grieve, & Hughes, 1978; Siegel et al., 1978; Winer, 1980). Misinterpretation of the question is the common feature in these proposals. That is, children interpret “more apples or more fruit” to mean “more apples or more other kinds of fruit,” and because there are only three pieces of non-apple fruit (oranges), they say there are more apples.

On the other hand Halford (1989) argued that many of the improved performances produced by alternative tests were no better than chance, or were amenable to alternative interpretations. Furthermore, techniques for estimating the number of answers attributable to misinterpretation or guessing have been developed (Hodkin, 1987; Thomas & Horton, 1997).

Alternative assessments have been devised, one based on a sorting task that was isomorphic to class inclusion but did not include potentially misleading questions (Halford & Leitch, 1989) and one based on property inference (Greene, 1994; Johnson, Scott, & Mervis, 1997). Understanding class inclusion entails recognition of the asymmetric relation between categories at different levels of the hierarchy. For example, properties of fruit apply to apples, but the reverse is not necessarily true, because apples may have properties not shared with other fruit. Halford, Andrews, & Jensen (2002) assessed category induction and class inclusion by equivalent methods, based on property inference. Relational complexity analysis showed that category induction is binary relational, because it entails a comparison of a class with its complement (e.g., birds and non-birds). Class inclusion is ternary relational because it necessarily entails an inclusive class (e.g., fruit), a subclass (e.g., apples), and a complementary subclass (non-apple fruit). When assessed by equivalent methods, class inclusion was found to be more difficult and performance on it was predicted by ternary relational tasks from other domains. This suggests that category induction and class inclusion are really the same paradigm at two levels of complexity.

Conservation, transitivity, and class inclusion are all ternary relational (Andrews & Halford, 1998; Andrews & Halford, 2002; Halford, Wilson, & Phillips, 1998) and this level of complexity is attainable by approximately twenty percent of four-year-olds, 50% of five-year-olds, 70% of 6-year-olds, and 78% of seven- and eight-year-olds. There is no age at which children suddenly attain all these concepts, as implied by some interpretations of Piagetian stage theory. Rather, the proportion of children who succeed increases according to a biological growth function. We can conclude
that all are attained at a median age of five years.

**Concept of Mind**

Children's ability to understand other people's mental states has been one of the most intensively researched topics in the past two decades. See Astington (1993) or Wellman, Cross, and Watson (2001) for reviews. Two main types of tasks have been employed—appearance-reality and false belief. Appearance-reality is tested by presenting children with an object that appears to be something else and asking them what it really is and what it appears to be. For example, Flavell, Green, and Flavell (1986) showed children a small white fish and then covered it with a blue filter and asked what color it was really and what color it appeared to their eyes. Children below about four years have difficulty recognizing both that the object is really white and that it appears blue. In a typical false-belief task (Wimmer & Perner, 1983), Person 1 hides an object in a box and leaves the room, then Person 2 shifts the object to a basket, and then Person 1 returns. Before age four, children have difficulty recognizing that Person 1 will look for the object in the box because he or she did not see it moved to the basket.

Numerous factors have been shown to influence children's concept of mind, including social-perceptual knowledge (Tager-Flusberg & Sullivan, 2000), understanding of mental states (Bretherton & Beeghly, 1982), and language (Austingon & Jenkins, 1999). Astington & Gopnik (1991) proposed a theory-theory, meaning that children's concepts of belief, desire, and pretence are linked in an explanatory framework. The more neutral term "concept of mind" is used here, even though "theory of mind" is in common use, because there are still doubts that children's understanding of the mind amounts to a theory. For example, telling children that people's thoughts can be wrong or reminding them of their own false beliefs does not raise three-year-olds' performance above chance. Leslie (1987) proposed an innate theory of mind mechanism, or module, that is specialized for processing social cues indicating mood, interest, or attention.

There is growing evidence that concept of mind is related to executive function (Carlson & Moses, 2001; Perner, Lang, & Kloo, 2002) and is partly a function of ability to deal with the appropriate level of complexity. Halford (1993; Halford, Wilson, & Phillips, 1998) analyzed the complexity of concept of mind tasks and showed it entails integrating three variables—the environmental cue, the setting condition, and the person's representation. Appearance-reality requires processing the relation between object color (white), the color of the filter (blue), and the percept (white or blue). Evidence that complexity is a factor in concept of mind has been produced by several groups of researchers (Andrews et al., 2003; Davis & Pratt, 1995; Frye, Zelazo, & Palfai, 1995; Gordon & Olson, 1998; Halford, 1993; Keenan, Olson, & Marini, 1998).

The analysis showing that concept of mind requires processing ternary relations suggests this should not be possible for chimpanzees because the most complex relation they have been shown to process is binary (Halford, Wilson, & Phillips, 1998). Although the issue has been controversial, a well-controlled study by Call and Tomasello (1999) tends to support this prediction. (See also Tomasello & Call, Chap. 25.)

**Scientific Thinking**

The topics reviewed so far on children's understanding of conservation, transitive inference, serial order, classification, cause, and biological processes are all important to the development of scientific and mathematical thinking. In this section, we consider some of the more advanced forms of scientific and mathematical reasoning in children. (See also Chap. 5 on causal reasoning by Buehner & Cheng, Chap. 23 on mathematical thinking by Gallistel & Gelman, and Chap. 29 on reasoning in science by Dunbar & Fugelsang.)

Whether children think as young scientists has been a major question of interest
motivated partly by evidence that categorization and concept of mind are driven by theories of the domain, as noted earlier. Kuhn et al. (1988) investigated how children assessed evidence in order to test hypotheses, and concluded that there was a considerable lack of scientific objectivity, especially among the younger children. Similarly, Klahr, Fay, and Dunbar (1993) found strong developmental effects in a study of ability to design experiments to determine rules underlying development of a robot. On the other hand, Ruffman et al. (1993) found evidence that six-year-olds have some understanding of how covariation evidence has implications for hypotheses about factors responsible for an event. There is also evidence that children as young as five recognize the evidential diversity principle—that we can be more confident of an induction from a set of diverse premises than from a set of similar premises (Heit & Hahn, 2001; Lo et al., 2002; Sloman, Chap. 3, this volume). A theoretical account of the development of inductive reasoning is given by Kuhn (2001), and a review of the development of scientific reasoning skills is provided by Zimmerman (2000).

**Time, Speed, Distance, and Area**

Understanding time, speed, and distance is interesting because it entails relations among three variables; speed = distance / time; and this relation should be accessible from other directions, so distance = speed × time, and so on. Matsuda (2001) found a progression from considering relations between two variables (e.g., between duration and distance or between distance and speed) at four years to integration of all three dimensions by age 11. Wilkening (1980) used an information integration theory approach in which the variance in children's judgments of distance was assessed as a function of speed and duration. In the information integration approach, reliance on a factor is indicated by a main effect, and reliance on the product of speed and duration is indicated by an interaction of these factors. Integration by an additive rule, speed + distance, is indicated by two main effects. Children from five years to adulthood showed evidence of the multiplicative rule. A similar assessment of children's understanding of area indicated gradual progression from additive rule (area = length + width) to a multiplicative rule (area = length × width) by adulthood. A cascade correlation model of time, distance, and velocity judgments is provided by Buckingham and Shultz (2000).

**Causal Reasoning**

Infants are able to perceive causal links between entities (Leslie & Keeble, 1987), but the causal reasoning of older children seems to be influenced by complexity (Brooks, Hanauer, & Frye, 2001; Frye et al., 1996) or by concept availability (Ackerman, Silver, & Glickman, 1990). The explanation may be that, as Leslie and Keeble (1987) suggest, the causal recognitions of infants are based on a modular process that is essentially perceptual. Modular processes are not typically influenced by cognitive complexity. The causal reasoning of older children probably depends on more conceptual or symbolic processes (Schlottman, 2001).

**Balance Scale**

Siegler (1981) applied the rule assessment approach to children's performance on the balance scale, yielding the four rules that were discussed in connection with McClelland's neural model (1995). Siegler's data showed Rule I (judgments based solely on weight) was used by five-year-olds, and they could also be taught to use Rule II (distance considered if weights are equal). Surber and Gzesh (1984) used an information integration approach and found that five-year-olds tended to favor the distance rule. Case (1985), Marini (1984), and Jansen and van der Maas (1997) generally supported Siegler's findings and saw little understanding of the balance scale before age five.

Relational complexity theory (Haldorf, 1993; Halford, Wilson, & Phillips, 1998) proposes that discrimination of weights with distance constant, or distances with weight constant, entails processing binary relations.
and should be possible for two-year olds. This prediction was contrary to previous theory (Case, 1985) and empirical observation (Siegler, 1981). Integration of weight and distance requires at least ternary relations and should emerge along with other ternary relational concepts at age five years. These predictions were confirmed by Halford et al. (2002).

**Concept of the Earth**

The development of children's concept of the earth (Hayes et al., 2003; Samarapungavan, Vosniadou, & Brewer, 1996; Vosniadou & Brewer, 1992) has special interest because it entails a conflict between the culturally transmitted conception that the Earth is a sphere and everyday experience that tends to make it appear flat. Resolution of this conflict entails recognition that the huge circumference of the earth makes it appear flat from the surface. There is also a conflict between gravity, naively considered as making objects fall down, and the notion that people can stand anywhere on the Earth's surface, including the southern hemisphere, which is conventionally regarded as "down under." This can be resolved by a concept of gravity as attraction between two masses, the Earth and the body (person) on the surface, but there is little basis for this concept in everyday life. The development of children's concept of the Earth provides an interesting study in the integration of complex relations into a coherent conception. Young children were found to attempt resolution of the conflicting ideas by, for example, drawing a circular earth with a horizontal platform inside for people to stand on, or as a flattened sphere to provide more standing room at the top. Nevertheless, there was a clear tendency for ideas to develop toward coherence.

**Conclusions and Future Directions**

Although acknowledging that predictions of future developments are inherently hazardous, it seems appropriate after an extensive assessment of the literature to try to identify some of the more promising developments. My bias is to look for developments that might provide a coherent body of theory because this is what the field of cognitive development, like the rest of psychology, needs most. I have identified four trends I feel deserve consideration in this respect.

**Neuroscience and the Biological Perspective**

The greatly increased knowledge of neuroscience and the use of brain imaging as a converging operation to help constrain theories of cognitive development represent major developments in the past two decades. Combined with the biological perspective, they do offer some hope of a coherent framework for viewing cognitive developmental data. The identification of changes in rates of neural and cognitive development is one example of what this field has achieved.

**Dynamic Systems**

Dynamic systems models have made considerable progress, and they are much more clearly linked to data than was the case a decade ago. They can provide new perspectives on important issues such as whether development is continuous or discontinuous or the fact that performance might be uneven across different indicators of the same task. The relative importance of different classes of observations might change fundamentally with this perspective.

**Transition Mechanisms**

Transition mechanisms and more advanced conceptions of learning have provided real conceptual advances and some of the most important empirical findings in the past decade. Neural net models have defined some potential mechanisms of concept acquisition that would almost certainly never have been recognized intuitively and would have been very difficult to discover using our contemporary empirical methods.
Analyses of Underlying Processes

Analyses of underlying processes using methods from general cognitive psychology and cognitive science can help bring order and clarity to the field. There are many examples of tasks that are superficially similar (such as transitive inference and transitivity of choice) yet entail fundamentally different processes, and it only creates confusion to categorize them together. Correspondingly, there are tasks that are superficially very different, yet may entail unimportant equivalences, unless we look beneath surface properties. Cognitive psychology has progressed to the point at which we can do this with reasonable confidence.

References


Bjorklund, D. F. (1997). In search of a metatheory for cognitive development (or Piaget is dead and I don’t feel so good myself). Child Development, 68, 144-148.

Boole, G. (1854/1951). An Investigation of the Laws of Thought, on which are founded the Mathematical Theories of Logic and Probabilities. New York: Dover.


Halford, G. S., Phillips, S., & Wilson, W. H. (under review). Structural complexity in
cognitive processes: The concept of representational rank.


