CHAPTER TWENTY-SEVEN

Information-Processing Models of Cognitive Development

Graeme S. Halford and Glenda Andrews

Information-processing models help us to understand the data of cognitive development, to gain new insights, and to frame new and useful questions. They also bring new concepts into the field from the related areas, including cognitive psychology, cognitive science, and neuroscience. These concepts tend to provide new ways of analyzing and understanding the processes of cognitive development. Specifically, information-processing theories have yielded measures of cognitive complexity, which can be a major organizing theme in cognitive development, if it is found that children tend to master more complex concepts later.

A dominant early approach has been the neo-Piagetian models, which sought to explain the course of cognitive development, as observed by Piaget and his collaborators, in terms of the growth of information processing.

Neo-Piagetian Models

Several models developed in parallel in this field, often with considerable interaction between the theorists.

McLaughlin

The model of McLaughlin (1963), although not the first information-processing theory in the field, can probably be regarded as beginning the modern era of models in this category. McLaughlin proposed that Piagetian stage was determined by the number of concepts that could be considered simultaneously. Piaget’s sensorimotor, preoperational, concrete operational and formal operational stages required $2^0 = 1$, $2^1 = 2$, $2^2 = 4$ and
2³ = 8 concepts to be considered simultaneously. To illustrate, consider a set of objects with two attributes – shape (triangle, non-triangle) and color (red, non-red). This defines four possible categories: red triangle, red non-triangle, non-red triangle, and non-red non-triangle. Thus two binary valued attributes yield four categories, or 2² = 4. Now consider a set of objects with three attributes – shape (triangle, non-triangle), color (red, non-red), and size (large, non-large). Now there are 8 possible categories of objects: large red triangle, large red non-triangle, … non-large non-red non-triangle. Thus three binary valued attributes yield 2³ = 8 categories.

McLaughlin proposed that the number of categories that a child could consider simultaneously was determined by memory span. A child with a span of 2 could consider 2¹ = 2 concepts, and this would put the child at the preoperational stage. A child with a span of 3 or 4 could consider, 2² = 4 concepts, and would be in the concrete operational stage. A child with a span or 5 to 8 could consider 2³ = 8 concepts, and would be in the formal operational stage. McLaughlin pointed to a correspondence between the development of span and the progression through Piaget’s stages of cognitive development.

However McLaughlin did not apply the theory to analyze or predict performance on any specific cognitive tasks. Nevertheless, McLaughlin’s suggestion inspired a lot of thought by subsequent cognitive developmentalists, including our own work. Interestingly, Feldman (2000) has shown that complexity of concept learning can be defined by the length of the shortest Boolean expression that is equivalent to the concept, an idea that is broadly consistent with McLaughlin’s. Feldman’s formulation applies to experimental concept learning (Shepard, Hovland, & Jenkins, 1961) whereas McLaughlin’s theory applies to cognitive development, but there is enough common ground here to suggest that this approach to analysis of complexity might not yet have run its full course. McLaughlin’s formulation was the first serious attempt to quantify complexity of concepts that are important in cognitive development.

Pascual-Leone

The model of Pascual-Leone (1970) was first theory of complexity in cognitive development to be empirically tested. He proposed that children’s cognitive functioning was governed by central-computing space, or M-space, that corresponded to the number of separate schemes that they could coordinate. The value of M was a + 1 at age 3 and increased by 1 every 2 years, reaching a value of a + 7 at age 15, where a is a parameter that represents the processing space required for instructions and for the general task situation, and which is constant over age. To illustrate, Pascual-Leone proposed that a child for whom M = a + 2 would be in the last substage of Piaget’s preoperational stage, and would on average be 5–6 years old.

Pascual-Leone developed the compound stimulus visual information (CSVI) task to assess M-space. The stimuli varied over eight dimensions, such as shape, color, size, whether the figure was closed or open, whether there was a circle in the centre, and so on. Children were trained to produce a specific response for the positive attribute on each dimension (e.g., raise a hand if the shape was square, clap hands if the color was red, etc.). Testing consisted of presenting stimuli varying in from five attributes
(for 5-year-olds) to eight attributes (for 11-year-olds) and determining the number of responses given out of the total for the stimulus (e.g., given a stimulus with five positive attributes, a child might give four of the five responses they had been trained to give for that set of attributes). The number of active schemes was then estimated using the Bose-Einstein occupancy model. A good fit to the expected age norms was obtained.

Pascual-Leone then demonstrated that the M-space demands of some well-known cognitive developmental tasks predicted the ages at which these were attained. For example, Pascual-Leone and Smith (1969) analyzed class inclusion as requiring an M-space of three, because it is a union of two classes, A and A′ that are included in B (e.g., apples and bananas are included in fruit). It therefore requires coordination of a scheme for each of these classes, in addition to the schemes representing instructions and the task situation. The requirement to coordinate three schemes is consistent with attainment of class inclusion at approximately age 7, when M-space reaches three schemes.

The conception which Pascual-Leone introduced in 1970 has resulted in a repertoire of techniques for assessing children’s processing capacity, and research related to the paradigm continues at the present time. A number of detailed comparisons have been made between Baddeley’s (1990) working-memory model and Pascual-Leone’s M-space (Baddeley & Hitch, 2000; de Ribaupierre & Bailleux, 1994, 2000; Kemps, De Rammelaere, & Desmet, 2000). See also a reply by Pascual-Leone (2000). The links between M capacity, inhibitory control, executive processes (e.g., shifting, updating), and processing speed have also been examined (Im-Bolter, Johnson, & Pascual-Leone, 2006; Johnson, Im-Bolter, & Pascual-Leone, 2003).

Case

Case (1985) proposed that children’s cognitive processes develop because they make better use of the available capacity. Specifically, Case proposed that total processing space (TPS) could be flexibly allocated to operating space (OS) or short-term storage space (STSS), i.e.:


\[ \text{TPS} = \text{OS} + \text{STSS} \]

Total processing space was held to be constant over age, after infancy. Demands for operating space declined with age because of increased processing efficiency, thereby making more of the TPS available as short-term storage space. The increase in short-term memory span with age was attributed to more efficient processing, which also became the main factor responsible for cognitive development, according to the model. Their empirical methods are well illustrated by Case, Kurland, and Goldberg (1982) who assumed that participants who processed faster were more efficient, and required less operating space, leaving more short-term storage space. Therefore their recall would be superior. Consistent with the theory, short-term memory span for words was found to be well predicted by speed of rehearsal. A striking finding was that when adults’ rehearsal rate was reduced to that of 6-year-olds their spans were reduced correspondingly.
These findings appeared to offer a straightforward and elegant solution to the problem of whether processing capacity increased with age. They were consistent with the claim by Chi (1978) that memory performance reflected familiarity or expertise rather than capacity. Chi measured recall of 10-year-olds’ and adults’ memories for digit strings and for chess positions. With digits, the usual finding that adults’ recall was superior was obtained. However the children in the sample had much greater knowledge of chess than the adults, and their recall of chess pieces on a board was superior to that of the adults. This finding, like that of Case et al. (1982) was taken to indicate that memory capacity was not responsible for increases in span with age. Note however that, while both studies demonstrate an effect of variables other than capacity, processing speed in the Case et al. study, and domain knowledge in Chi’s study, they do not assess capacity. Neither Case et al.’s nor Chi’s (1978) findings rule out an increase in capacity with age, which is what more recent evidence indicates (Andrews, Halford, Bunch, Bowden, & Jones, 2003; Cowan, 2001; Halford, Cowan, & Andrews, 2007; Quartz & Sejnowski, 1997).

The theory that processing space is a single resource that can be flexibly allocated to processing or storage is hard to reconcile with findings from working-memory research. Halford, Maybery, O’Hare, and Grant (1994) gave children a memory preload, comprised of a string of digits, then a cognitive task to perform, then they were asked to recall the preload. This procedure separates the processing and storage demands that are combined in many working-memory measures, and the difficulty of the cognitive task was manipulated independently of the concurrent memory-storage task. They found that a concurrent short-term memory load does not interfere with reasoning as would be expected according to Case’s position.

Another aspect of Case’s early work has withstood the test of time. This is the insight that cognitive development depends heavily on learning to use the available capacity more efficiently. Processing capacity is limited in both children and adults (Cowan, 2001; Halford, Cowan, & Andrews, 2007; Luck & Vogel, 1997) and therefore cognitive processes require strategies that utilize the limited capacity effectively. The postulate that total processing capacity remains constant over age is no longer tenable, but it is true that increased efficiency in the way capacity is utilized accounts for a lot of what develops. The historic contribution from this phase of Case’s work may be that he was the first to realize the importance of processing efficiency in cognitive development, and to build this principle into a systematic theory.

He also anticipated more recent neuroscience evidence on growth of processing capacity. Case (1992b) reviewed evidence that the frontal cortex continues to develop for at least two decades after birth, and electroencephalographic (EEG) coherence between frontal and posterior lobes also increases. He then drew attention to correspondence between growth of EEG coherence and development of working memory over ages 4–10 years. This work arguably marks the beginning of attempts to base systematic cognitive development theory on neuroscience data.

In his later work Case (1992a, 1992b; Case et al., 1996) developed the concept of central conceptual structures, defined as “an internal network of concepts and conceptual relations that plays a central role in permitting children to think about a wide range of
situations at a new epistemic level, and to develop a new set of control structures for dealing with them” (Case, 1992a, p. 130). Development of central conceptual structures passes through the sequence of four major neo-Piagetian stages:

The sensorimotor stage is concerned with coordination between actions and reactions, or between actions and effects, for example when pushing a lever on a piece of apparatus sets it vibrating, producing an interesting sight or sound.

The interrelational stage is concerned with relations, such as one end of a beam moving down while the other moves up. At more advanced levels, coordination between relations occurs, so placing a heavier weight on one end of a beam and a lighter weight on the other is linked to one end of the beam going down while the other end goes up.

The dimensional stage is concerned with dimensional units, such as numbers, weights, or distances. Thus the child has progressed from relations between objects or events to thinking about relations between dimensional units such as number of weights, or between distances.

The vectorial (or abstract dimensional) stage is one in which dimensional structures are coordinated to produce abstract systems of dimensions, for example when ratios of weights and distances are converted to new ratios having a common term.

Each major stage is divided into substages that represent increasing structural complexity within the major stage. Complexity is defined by the number of nested goals required to perform the task. In order to predict which side of a balance beam would go down, at the simplest, preliminary substage there is a single goal, to determine which side of the balance would go down. The strategy is to simply look at the balance beam to determine which side looks heavier. At the next, unifocal substage, there is a hierarchy of two nested goals, to predict which side of the beam will go down, and a subgoal to determine which side has the larger number of weights. At the bifocal substage there are three nested goals, and at the elaborated coordination substage there are four nested goals.

The number of goals children can maintain, and hence the complexity of the problems they can solve, is determined by the size of STSS, which increases from 1 to 4, due to both maturation and experience. This progression occurs within each of the major stages. This means that the short-term storage load is reset to 1 when the child progresses to the next major stage, so a task that imposed a load of 4 at the elaborated coordination substage of interrelational stage would impose a load of 1 at substage 1 of the dimensional stage. The transition from one stage to another is achieved by coordinating two existing structures into a higher-order structure.

Case (1992a; Case et al., 1996) applied the theory to a wide range of domains, including scientific and mathematical knowledge, spatial and musical reasoning, understanding narrative and social roles, and motor development. However the metric is not general to all tasks, but is specific to performance within one of the major stages.

Fischer

The theory of Fischer (1980; Fischer & Rose, 1996) was based on cognitive skill theory, where skill refers to control over sources of variation in a person’s own behavior. There
are four major stages or tiers: the reflex, sensorimotor, representational, and abstract. Within each tier there is a recurring cycle of four levels: the set, mapping (of sets), system (composition of mappings), and system of systems. As with Case’s theory, the highest level of one tier is shared with the lowest level of the next, and represents a transition between tiers.

A set is a source of variation over which some cognitive process exercises control. In an action, the person can control the relevant variations in the behaviors on things. An infant who can consistently grasp a rattle has a set for grasping that rattle. (Fischer, 1980, p. 481)

In more conventional terms, a set is really a variable, but it implies correspondences between events or objects and actions: “The thing is always included with the behavior in the definition of a set” (Fischer, 1980, p. 481).

A set is the lowest level of structure within a tier. A single sensorimotor set would typically develop around 15–17 weeks of age. The next level is a mapping between sets. A sensorimotor mapping would typically develop at 7–8 months. An example would be coordination of looking at an object in order to grasp it. The next level is a system, which is really a coordination of two mappings. An example would be when the child drops a piece of bread and watches it fall, then breaks off a crumb and watches it fall. There is a mapping between dropping and watching the piece of bread and another between dropping and watching the crumb. By varying one set, the dropping, and observing the result, seeing the bread fall, the child develops a concept of means–end links. The next level is a system of sensorimotor systems, and is a mapping between two systems from the previous level. However it is also the first level of the representational tier, and is a single representational set.

Fischer (1980) illustrates the levels of the representational tier with a spring-and-cord gadget, in which a weight hangs from a cord that passes around a pulley to a coil spring that is attached to a vertical surface. As extra weight is added the spring stretches and the string moves around the pulley so that the horizontal part of the cord becomes shorter and the vertical part becomes longer, though the overall length of the cord is of course constant, which is a form of conservation of length (Piaget, 1950). A 4- to 5-year-old child who had experience with the gadget could use size of the weight to control the length of the spring. This is mapping one set, (variable) weight, into the other, spring length. The child might also make mappings between horizontal to vertical length of cord, or between any two of the four sets: weight, vertical and horizontal cord length, and spring length. However, understanding is disjointed because each of these mappings exists independently, and there is no integration such as recognizing that the horizontal and vertical lengths of cord are segments of one cord of constant length. A representational system is formed by relating two representational mappings. An example would be a mapping between the horizontal and vertical lengths of cord with one weight, and another mapping between the horizontal and vertical lengths with a different weight. Relating these mappings shows how the horizontal length decreases as the vertical length increases, and vice versa, which in turn leads to recognition that changes in the lengths compensate each other, leading to the idea that
the total length of the string is conserved. Thus the subtle and important concept of conservation emerges from the coordination of lower-level structures. The next level is a system of representational systems, which is a single abstract set. Development then proceeds through the remaining three levels of the abstract tier: the abstract mapping, abstract system, and the system of abstract systems, which is identified with principles.

Complexity increases within a tier in a manner that bears a striking correspondence to McLaughlin’s (1963) complexity scale. The four levels within a tier comprise a set, then a mapping between two sets, then a system which is a mapping between two mappings of sets and is equivalent to four sets, then a mapping between two systems, which comprises eight sets (see, for example, Fischer, 1980, figure 2, p. 490). This corresponds to McLaughlin’s (1963) four levels defined as $2^0 = 1$, $2^1 = 2$, $2^2 = 4$ and $2^3 = 8$ concepts considered simultaneously. This again illustrates the underlying common ground between different information-processing theories of cognitive development.

Fischer also has an extensive consideration of transformation rules for creating the transition from one level to another. The first rule is intercoordination, which is a process of combining skills at one level to produce a skill at the next level. An example would be the two mappings between horizontal and vertical lengths of cord under different weights, as mentioned earlier. The mappings are intercoordinated to create a system, and recognition of the compensating changes between horizontal and vertical lengths emerges from the intercoordination. The other rules are compounding, focusing, substitution, and differentiation.

Fischer’s (1980) formulation has been retained in essence in his later work, but there have been two major developments of the model. One has been to link the model to dynamic growth functions (Van der Maas & Molenaar, 1992; van Geert, 1993) while the other has been to link it to spurts in brain growth (Fischer & Bidell, 1998; Fischer & Rose, 1996; Thatcher, 1994). Dynamic systems models of cognitive development are considered by Thelen and Smith (2006) but we will try to indicate briefly how Fischer’s theory has developed along these lines. Firstly, it is proposed that the major reorganizations between levels, as outlined above, correspond to growth spurts or other discontinuities in brain growth. Furthermore, these dynamic changes can occur concurrently in many independent systems, which might be localized in different regions of the brain. It is also proposed that each new level is marked by a new behavioral control system, which is supported by a new kind of neural network (Fischer & Rose, 1996). Fischer, Stewart, and Stein (2008) propose that knowledge is built from repeated reconstructions that move towards higher complexity and abstraction, though with many reversals and recoveries. Fischer also proposes that the recurring cycles of development that occur in each tier are supported by observations of brain growth. A more comprehensive summary is given by Fischer and Bidell (2006).

The implications of the recent developments of Fischer’s theory are profound and sophisticated. They offer a resolution of the anomaly, crucial to understanding development, that there is both variability and consistency in cognitive development. Dynamic systems produce variability from a relatively small set of common processes. The application of dynamic systems theory to the database provided by cognitive developmental stage theory is one of the greatest achievements of Fischer and his colleagues.
Attempts at Synthesis

Attempts have been made to synthesize the neo-Piagetian theories (Chapman, 1987, 1990; Chapman & Lindenberger, 1989; Demetriou, Doise, & van Lieshout, 1998; Demetriou, Efklides, & Platsidou, 1993). Chapman’s approach was arguably more directly based on information-processing concepts. He postulated 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. (Chapman & Lindenberger, 1989, p. 238)

We will consider his analysis of class inclusion. A concrete example might comprise a set of beads, some red (A) and some blue (A’), all of which are wooden (B). Therefore A and A’ are included in B (A ∪ A’ = B). The child is asked to decide whether there are more red than wooden beads. Solving the problem entails assigning values to the class variables, A, A’, and B, that is recognizing that A = red beads, A’ = blue beads, and B = wooden beads, which is a form of variable binding. Chapman and Lindenberger refer to this as the coordination of intension (the defining properties of the class) and extension (objects belonging to the class).

Chapman (1987) analyzed tasks in terms of the number of schemes required for solution:

One of the main hypotheses generated by the proposed model is that the notion of attentional capacity can be explicated in terms of the number of representational schemes coordinated by an inferential scheme. (Chapman, 1987, p. 310)

However he interpreted schemes in a way that was very like variables:

A further property of representational schemes is that they can be generalized beyond their immediate context and embedded in more abstract schemes without losing their identity. (Chapman, 1987, p. 309)

Chapman was very much aware of the inherent difficulties in objective analyses of complexity. After a penetrating review of the neo-Piagetian theories of Case, Pascual-Leone, and Halford, Chapman (1990) commented:

In summary, neo-Piagetian theorists have not yet developed a method of task analysis (a) that is sufficiently rigorous to result in unambiguous predictions and (b) that can be applied with equal facility to both cognitive tasks and measurement tasks. (Chapman, 1990, p. 273)

In his later work, Chapman was more explicit in analyzing complexity in terms of variables. Perhaps therefore one of the most important achievements of Chapman and his collaborators was to have realized that the best way to analyze complexity of cognitive tasks is to determine the number of variables that have to be instantiated in parallel.
Variables can be analyzed more objectively than schemes. The concept of scheme is so flexible that it may be difficult for independent observers to agree on the number of schemes required for a task.

**Processing Speed**

Evidence for a global processing-speed factor that increases with development was provided by Kail (1986, 1988a, 1991; Kail & Park, 1992). The methodology was based on measuring changes in processing speed with age across a number of rather different tasks, and is well illustrated by Kail (1988a). Children were tested on a memory search task in which they studied a set of one, three, or five digits, then a single-digit probe was presented and children had to decide whether it had been in the study set. Visual search was similar except that the study set was a single digit, and children had to determine whether this digit appeared in a probe set of one to five digits. Processing speed can be measured in both tasks by the slope of the function relating probe set size to response time. Other tasks used included mental rotation, in which children had to judge whether a pair of letters presented in different orientations were identical or were mirror images, and a mental addition task in which children had to determine the correctness of sums such as $3 + 8 = 10$. There was also an analogical reasoning task with two $3 \times 3$ matrices. Children had to determine whether the geometric figures in the cells changed according to the same rule in both matrices. It was found that the change in processing speed over the age range 8–22 years was very similar across all task domains, and was best fitted by an exponential function that was common to the tasks. This led Kail to propose that changes in processing speed over age reflected processing capacity rather than learning.

While these findings offer some of the strongest evidence for growth of processing capacity with age, they have not been without controversy. One issue has been whether the findings are compatible with some kind of learning model (e.g. Stigler, Nusbaum, & Chalip, 1988; but see also Kail, 1988b, 1990, 1991). The global nature of the processing-speed factor has also been challenged by Ridderinkhof and van der Molen (1997) who argued that process-specific factors may be involved. There is also a cause-and-effect question: Does increased speed cause an increase in processing capacity, or the reverse? A possible answer comes from neural net models. Increased capacity reduces processing time in some neural nets because it reduces the number of cycles required for the net to settle into the solution that best fits the parallel-acting constraints. Alternatively it is possible that higher processing speed permits more information to be processed before activation decays. This issue seems likely to remain active for some time (see, for example, Cowan, 1998; Halford, Wilson, & Phillips, 1998a). These controversies notwithstanding, the discovery of a global processing-speed factor must be considered one of the major achievements in the field. Processing speed has emerged as a major factor in cognitive development across the lifespan (Cerella & Hale, 1994; Kail & Salhouse, 1994; Salhouse, 1996) and it has found application in a number of other contexts (Kail, 1997, 1998, 2000; Kail & Hall, 1999).
Further issues have been raised by Anderson (1992), who proposes that individual differences reflect processing efficiency, whereas cognitive development depends on knowledge elaboration and maturation of modules. He reviews evidence that individual differences in intelligence are stable over the course of development, and are related to measures of processing speed such as inspection time, which is the exposure duration required to detect which of two vertical lines is longer. This has been found to correlate with intelligence (Nettelbeck, 1987).

Cognitive development, on the other hand, is considered to be heavily dependent on maturation of specialized modules that are functionally independent, complex processes of evolutionary importance, independent of general intelligence. Anderson proposes that there are modules for perception of 3-dimensional space, phonological encoding, syntactic parsing, and theory of mind. The mechanisms that enable us to see in 3-dimensional space are specialized for processing visual information, and are complex, but do not correlate with intelligence. Although it is likely that some functions, especially in perception, are modular, it is a much greater leap to propose that cognitive development and individual differences in intelligence exist in separate, watertight compartments. It seems more likely that cognitive development also depends on acquisition of domain-general processes, such as memory storage and retrieval functions, and analogical reasoning (Halford & Andrews, 2004, 2007).

Cognitive Complexity

The orderly interpretation of findings in cognitive development depends on having a metric for cognitive complexity. It is only by comparing complexities that we can determine whether young children's performance on a given task is precocious.

Relational Complexity theory

One such approach, which has its origins partly in the neo-Piagetian approach, is Relational Complexity (RC) theory (Halford et al., 1998b; Halford, Cowan, & Andrews, 2007). Complexity is defined in terms of relations that can be processed in parallel. The essential idea is that each argument of a relation represents a source of variation, or a dimension, and an \( n \)-ary relation is a set of points in \( n \)-dimensional space. Thus relations of higher \emph{arity} are more complex, so a unary relation is less complex than a binary relation, which is less complex than a ternary relation, and so on. Processing load increases with relational complexity, and empirical evidence indicates that quaternary relations are the most complex that adults can process in parallel (Halford, Baker, McCredden, & Bain, 2005). This is consistent with limitations in short-term memory capacity (Cowan, 2001). Normative data suggests that unary relations are processed at a median age of 1 year, binary relations at 2 years, ternary relations at 5 years, and quaternary relations at 11 years.
Concepts too complex to be processed in parallel are handled by *segmentation* (decomposition into smaller segments that can be processed serially) and *conceptual chunking* (recoding into less complex relations, but at the cost of making some relations inaccessible). For example, \( v = \frac{st}{t} \) (velocity = distance/time) is a ternary relation, but can be recoded to a unary relation, a binding between a variable and a constant. However this makes relations between velocity, distance, and time inaccessible (e.g., if velocity is represented as a single variable, we cannot answer questions such as “How is speed affected if the distance is doubled and time held constant?”).

Complex tasks are normally segmented into steps, each of which is of sufficiently low relational complexity to be processed in parallel. The steps are processed serially. Expertise is important for devising strategies that reduce the complexity of relations that have to be processed in parallel, though a lower limit is usually imposed by the structure of the task. The effective relational complexity of a task is the most complex relation that has to be performed in parallel, using the most efficient strategy available. Complexity analyses are based on principles that are common across domains and methodologies.

We can illustrate relational complexity with the class inclusion task, discussed earlier, comprised of wooden beads, most of which are red, while the remainder are blue. Children are asked “Are there more wooden beads or more red ones?” Young children tend to say there are more red ones. The possible causes of error have been the subject of much controversy (Breslow, 1981; Bryant & Trabasso, 1971; Halford, 1993; Hodkin, 1987; Markovits, Dumas, & Malfait, 1995; McGarrigle, Grieve, & Hughes, 1978; Pears & Bryant, 1990; Siegel, McCabe, Brand, & Matthews, 1978; Thayer & Collyer, 1978) but when allowance is made for these, we still have a complexity factor that influences performance. The problem is that in order to determine which class is the superordinate and which are the subclasses, children must consider the relations among the superordinate class and the two subclasses. A class such as wooden beads is not inherently a superordinate, and its status is defined by its relations to the subordinates. That is, wooden is a superordinate because it includes red and blue beads. All three sets and the relations between them are necessary to understand that the subclasses are included in the superordinate class. This entails a ternary relation. If the task is represented as a series of separate binary relations that are not integrated, then the full implications of the entire relational structure among the classes (e.g., that the superordinate class is necessarily more numerous than the major subclass) will be missed.

Andrews and Halford (2002) investigated the emergence of ternary-relational processing in the class inclusion task as well as five other content domains (transitive inference, hierarchical classification, sentence comprehension, cardinality, and hypothesis testing). The percentages of children who processed ternary relations were 16% at age 3–4 years, 48% at age 5, 70% at age 6, and 78% at age 7–8. The majority of children at each age succeeded on comparable binary-relational items. In this and other research (e.g., Andrews & Halford, 1998; Bunch, Andrews, & Halford, 2007; Halford, Andrews, & Jensen, 2002; Halford, Bunch, & McCredden, 2007) relational complexity was manipulated independently of other factors, and substantial complexity effects were observed. An important point about capacity to process complex concepts is that it develops in parallel with the acquisition and organization of knowledge, and there is considerable interaction
between the two sets of processes. Knowledge and complexity aspects of cognitive development have been considered in detail by Halford, Cowan, and Andrews (2007), who also provide objective principles for analysis of cognitive complexity.

A number of unequivocal developmental predictions have been made in advance using RC theory (Halford, 1993). For example, 2-year-olds should be able to discriminate either weight or distance, but not both, on the balance scale, which has been confirmed by Halford, Andrews, Dalton, Boag, and Zielinski (2002). It was also predicted that structural complexity would be a factor in concept of mind (Halford, 1993), which was confirmed by Davis and Pratt (1995), Frye, Zelazo, and Palfai (1995), Gordon and Olson (1998), Andrews et al. (2003), and Keenan, Olson, and Marini (1998).

**Cognitive Complexity and Control theory**

Another complexity approach to cognitive development is Cognitive Complexity and Control (CCC) theory (Zelazo, Müller, Frye, & Marcovitch, 2003). This theory, which developed independently of RC theory, focuses on the complexity of the rules children are able to use. There are four types of rules that vary in complexity. The simplest are single stimulus–reward associations, which can be represented either implicitly or explicitly (e.g., when reversal of the association is required). Condition–action (if–then) rules are more complex than associations. In univalent rule pairs each stimulus is associated with a separate response (e.g., red light – stop; green light – go). Bivalent rule pairs are more complex because each stimulus is associated with two different responses, the correct response being determined by the context. For example, in the Dimensional Change Card Sorting (DCCS; Zelazo, 2006) task, a red boat is sorted with a red flower in the color game, but when the context changes to the shape game, the red boat is sorted with the blue boat. Higher-order rules integrate pairs of bivalent rules into rule hierarchies which facilitate selection among task sets (Zelazo et al., 2003). Older children construct more complex rules than younger children. Reversal of stimulus–reward associations has been demonstrated in non-human animals and children as young as 30 months (Overman, Bachevalier, Schuhmann, & Ryan, 1996). From 3 years, children can use a pair of arbitrary univalent rules, but they experience difficulty with bivalent rule pairs. Around 5 years, children can integrate two incompatible pairs of bivalent rules into a single rule system via a higher-order rule (Zelazo, Jacques, Burack, & Frye, 2002). Predictions derived from CCC theory have been supported by empirical research in that success on tasks that involve higher-order rules is not usually observed before 4.5 to 5 years of age (e.g., Frye et al., 1995; Kerr & Zelazo, 2004).

CCC-R (i.e., Cognitive Complexity and Control theory, revised; see Zelazo & Müller, chapter 22, this volume) and RC theories developed independently, but they share common ground. Analyses based on rule complexity and relational complexity are translatable in many situations. For example, according to CCC theory, correct responding on the DCCS (Zelazo et al., 2003), false-belief and appearance–reality tasks (Frye et al., 1995), and the Children’s Gambling Task (Kerr & Zelazo, 2004) requires higher-order rules. In RC theory, these tasks are ternary-relational (Andrews et al., 2003; Bunch et al., 2007; Halford, Bunch, & McCredden, 2007). However, whereas CCC theory
applies to tasks with a hierarchical structure, RC theory can also be applied to tasks with non-hierarchical structures.

Both complexity approaches propose that maturation of the frontal lobes underpins age-related increases in children’s ability to deal with increasing complexity. There is abundant evidence for the protracted maturation of prefrontal regions. Measures of synaptic density and elimination (Huttenlocher & Dabholkar, 1997), and of myelination (Paterson, Heim, Friedman, Choudhury, & Benasich, 2006) show that prefrontal regions are the last to reach maturation. Dramatic changes in myelination continue in the dorsal, medial, and lateral regions of the frontal cortex during adolescence (Nelson, Thomas, & De Haan, 2006). Glucose metabolism in the frontal and association cortices increases between 8 and 12 months (Chugani & Phelps, 1986). Within the frontal lobes grey matter maturation occurs earliest in the orbitofrontal cortex (OFC), Brodmann area (BA) 11, later in the ventrolateral prefrontal cortex (VL-PFC; BA44, BA45, BA47), and later still in dorsolateral prefrontal cortex (BA9 and BA46), coinciding with its later myelination (Gogtay et al., 2004).

Studies of brain function in adults show that relational processing involves the prefrontal cortex. Waltz et al. (1999) studied patients with brain lesions due to dementia. Patients with damage to the dorsolateral prefrontal cortex (DL-PFC) could use single binary-relational premises but they could not integrate two premises in a transitive inference task. Premise integration is ternary-relational (Andrews & Halford, 1998; Halford et al., 1998b). Similar results were obtained for a matrix reasoning task that required relational integration. Brain-imaging studies have implicated the DL-PFC and rostrolateral prefrontal cortex (RL-PFC) in relational integration, and as relational complexity increases more anterior regions appear to be recruited (Christoff & Owen, 2006; Christoff et al., 2001; Kroger et al., 2002). Smith, Keramatian, and Christoff (2007) demonstrated that a relational match-to-sample task reliably activates RL-PFC in adults. Ramnani & Owen (2004) proposed that the functions of the anterior prefrontal cortex (APFC) are distinguished more by the coordination and integration of information processing than by cognitive domain. This is consistent with Christoff and Owen's (2006) proposal that the functions of the RL-PFC are related more to cognitive complexity than to a cognitive domain.

Brain-imaging studies of the developing brain are still quite scarce. However performance on the Piagetian A-not-B task appears to be associated with maturity of brain systems involving the prefrontal cortex. Bell and Fox (1992) found that toleration of longer delays in 12-month-olds was associated with more mature patterns of EEG brain activity. Bell (2001) recorded EEG activity while 8-month-olds performed a looking version of the A-not-B task. High-performing infants showed task-related increases in EEG power in four scalp regions (frontal pole, medial frontal, parietal, occipital), suggesting the involvement of both frontal and non-frontal brain regions. They also showed increased EEG coherence between medial frontal and parietal sites, suggesting that these regions were working together in the task, and lower coherence between two frontal pairs of electrodes in the right than left hemisphere, consistent with more advanced differentiation. Low-performing infants showed more hemispheric symmetry. Lesion research with non-human primates implicates the DL-PFC (BA8, BA9, BA10), but not the hippocampus in A-not-B task performance (Diamond, 1990; Diamond, Zola-Morgan, & Squire, 1989).
Bunge and Zelazo (2006) proposed a brain-based account of the development of rule use in childhood which links the four rule types to different prefrontal cortex sub-regions. Simple stimulus–reward associations involve the OFC (BA11). Univalent and bivalent conditional rules involve the VL-PFC (BA44, BA45, BA47) and the DL-PFC (BA9, BA46). Higher-order rules recruit the RL-PFC (the lateral portion of BA10).

Imaging studies involving children aged from 8 to 12 years (Crone et al., 2009) are consistent with these complexity accounts. Extending brain-imaging research to examine acquisition of complex rules and relations in younger 3- to 8-year-old children is an important area for future research, although the difficulties associated with using these techniques with young children are far from trivial.

Levels of Cognitive Functioning

Levels of cognitive function have been defined by a number of writers, including Campbell and Bickhard (1986) and Karmiloff-Smith and her collaborators (Clark & Karmiloff-Smith, 1993; Karmiloff-Smith, 1992, 1994). The most influential at present is probably that of Karmiloff-Smith, based on the implicit–explicit distinction. It is considered by Westermann, Thomas, and Karmiloff-Smith (chapter 28, this volume), but has links to the other information-processing approaches. Level-I, or implicit knowledge, is an effective basis for performance, but is not accessible to other cognitive processes, and cannot be modified strategically. It is knowledge in the system but not knowledge to the system. There are three levels of explicit knowledge that become progressively more accessible, modifiable, and available to consciousness and verbal report. Unlike neo-Piagetian models, representational redescription occurs independently in each domain, and is not age- or stage-linked. Nevertheless representational redescription attempts to integrate Piagetian constructivism and Fodor’s (1983) concept of innate knowledge.

The nature of representational redescription has been only partly specified. Phillips, Halford, and Wilson (1995) proposed that the implicit–explicit distinction can be captured by the distinction between associative and relational knowledge. Karmiloff-Smith (1994) has speculated about neural net models that might make the transition from one level to another. However much remains to be done to address this fundamental issue.

Process Models of Cognitive Development

A number of information-processing models of concept acquisition have been developed. 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 (e.g. by spacing objects more widely, increasing the length of the row, without
adding any items) then the child is asked whether each row still contains the same number, and whether they are the same or different.

The preconserving child cannot answer this question. This is represented in Q-SOAR as an impasse. The model then searches for a solution to the problem, using a procedure based on the work of Klahr and Wallace (1976). This entails quantifying the sets before and after the transformation, noting that they were the same before the transformation, that they are the same after the transformation, and finally that the transformation did not change the relation between the sets. This process uses knowledge already available, including quantifying the sets, comparing them, recalling results of quantification and comparison, and noticing the effect of transformations (both conserving and non-conserving). With repeated experience, the model gradually learns to classify the action of spacing out the items as a conserving transformation.

Following work on rule assessment (Briars & Siegler, 1984; Siegler, 1981), Siegler and his collaborators conducted an extensive study of strategy development (Siegler, 1999; Siegler & Chen, 1998; Siegler & Jenkins, 1989; 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 choose 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 3 and a set of 5, perhaps using fingers), and the min strategy of counting on from the larger set (e.g., 4, 5, 6, 7, 8, so 3 + 5 = 8).

Their early strategy-choice model was based on distribution of associations (Siegler & Shrager, 1984). The idea is that each addition sum is associated with answers of varying strengths 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”). This model 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. 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 where 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 depending on 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 over age.
Relational Knowledge and Analogy

One of the most fundamental problems in cognitive development is children’s acquisition of relational and dimensional knowledge, yet there has been little systematic study of it. One of the most interesting research projects in this area is that by Smith, Gasser, and Sandhofer (1997). Examining evidence from children’s word acquisition, Smith noted that dimensional adjectives (e.g., “wet,” “soft,” “big,” “red”) are learned relatively slowly as compared with, say, nouns. Smith postulated that to learn dimension words, children must learn three kinds of mappings; between words and objects (“red” for red objects), word–word maps (“red,” “blue,” etc., are associated with color), and property–property maps (“They are the same color”). Smith argues that early use of relational terms is holistic, but relational terms gradually become organized into dimensions (by about 5 years of age). This enables recognition that “more than” is the opposite of “less than” and recognition of transitivity of relations (e.g. $a > b$ and $b > c \rightarrow a > c$).

Recognition of the role of analogy in cognition and cognitive development increased rapidly in the 1980s. Gentner (1983) provided a workable conceptualization of human analogical reasoning with her theory that analogy is a mapping from a familiar structure, the base, to an unfamiliar structure, the target. The mapping is validated by correspondence between the structures in base and target. One difficulty is that analogical reasoning has sometimes been difficult to produce in the laboratory (Gick & Holyoak, 1983) but it occurs readily in real life (Dunbar, 2001) so this difficulty might be overcome by using more naturalistic procedures. Computational models of analogical mapping (e.g., Falkenhainer, Forbus, & Gentner, 1989; Holyoak & Thagard, 1989; Hummel & Holyoak, 1997) helped to clarify the nature of the process, and its role as a mechanism of cognitive development has been explored (e.g. Halford, 1993). The DORA (Discovery Of Relations by Analogy) model (Doumas, Hummel, & Sandhofer, 2008) has also provided an existence proof for acquisition of structured knowledge by self-supervised learning from examples.

Much of the developmental interest in analogy has centered around the age of attainment (Gentner & Ratterman, 1991; Gentner, Rattermann, Markman, & Kotovsky, 1995; Goswami, 1996; Goswami & Brown, 1989). There is no reason to doubt that simple proportional analogies of the form A is to B as C is to D can be performed by children under 5 years. Indeed, according to complexity analyses by Halford and his collaborators (Halford, 1993; Halford et al., 1998b) they should be possible at around 1.5–2 years. Analogies that require parallel processing of more complex relations are predicted to be consistent with the age norms for that level of relational complexity.

Symbolic Neural Net Models

Neural net models, considered by Westermann, Thomas, and Karmiloff-Smith (chapter 28, this volume), have become one of the most important types of computational models of cognitive development. However questions about the ability of some early models to
account for symbolic processes (Fodor & Pylyshyn, 1988; Marcus, 1998a, 1998b; Smolensky, 1988) have given rise to symbolic neural net models (Halford, Wilson, et al., 1994; Hummel & Holyoak, 1997; Shastri & Ajjanagadde, 1993). We will briefly consider two attempts to build symbolic neural net models of cognitive development.

The first is a neural net implementation of the processing complexity theory of Halford et al. (1998b) discussed earlier. It is based on the Structured Tensor Analogical Reasoning (STAR) model of Halford, Wilson, et al. (1994). The essential idea is that the structural properties of higher cognition can be captured by the representation and processing of relations. Representation of a relation entails a symbol for the relation, plus a representation of the related entities. These must be bound together in a way that preserves the truth of the relation. Consider a simple binary relation, such as larger-than. An instance of larger-than, in predicate calculus notation, is larger-than(elephant,mouse). In our model we represent this by having a set of units representing each component, larger-than, elephant, and mouse. Each set of units corresponds to a set of activation values, or vector. The binding is represented by computing the tensor (outer) product of the vectors, producing a 3-dimensional matrix. More complex relations correspond to binding more entities, and therefore to tensor products of higher rank. Thus a binary relation entails binding three entities, the symbol and two arguments, a ternary relation to binding four entities, the symbol and three arguments, and so on. Tensor product nets are shown schematically in figure 27.1, together with the approximate Piagetian stage with which each level of relational complexity is identified in the theory of Halford et al. (1998b).

This model is designed to represent structure with a neural net architecture. Thus the components retain their identity in the compound representation. In the Rank 3 tensor product representing larger(elephant,mouse) the vectors representing larger, elephant, and mouse are retained. Also, any component can be retrieved, given the remaining components, a property that Halford et al. (1998b) call omni-directional access. The relations in a net can be modified online, by changing the relation symbol. This is one of the properties of explicit knowledge defined by Clark and Karmiloff-Smith (1993). The model can handle higher cognitive processes such as analogical reasoning, and mathematical operations. Because the number of neural units, and therefore the computational cost, increases exponentially with complexity of relations, the model provides a natural explanation for the link between processing loads and relational complexity observed by Halford et al. (1998b). On the other hand this model does not have the learning functions that are a major benefit of multilayered net models, such as the balance-scale model of McClelland (1995), and consequently does not handle the emergent properties of these models. Therefore at the present time it appears that symbolic and multilayered net models should be seen as complementary, and the next step forward might depend on hybrids, or on models that capture the properties of both classes of nets.

The model of Smith, Gasser and Sandhofer (1997) simulates children’s acquisition of dimensional terms. It employs the three-layered net architecture, but can probably be categorized as a symbolic model because it specifically addresses acquisition of symbolic knowledge. An object with four perceptible attributes is coded in the input layer. There is a separate input that codes the relevant dimension, such as color. There is a further input indicating whether two successive objects are the same or different on the specified dimension. The representations in the hidden layer are copied to a perceptual buffer
The net is trained to produce the appropriate dimensional attribute (e.g., red). The training is constrained by property–property maps. If two successive objects are the same on the specified dimension, the output must be the same for both of them. The success of the model tends to support the authors' claim that culturally transmitted information about the sameness of objects on specified dimensions (e.g., “these flowers are the same color”) is important for children's acquisition of dimensional terms.

**Conclusion**

One of the most striking things about research on information-processing approaches to cognitive development is the richness of both empirical and theoretical work in the field.
Another is that information-processing approaches to cognitive development have told us a lot about the nature of the underlying processes in cognitive development. This means that we now have a lot more information about what is happening when a child is performing a task. This yields a lot of insights and offers genuinely new ways of understanding many issues. A third observation is the success with which information-processing conceptions have been linked to neuroscience and to neural net models. While no one would pretend that all problems have been solved in these areas, they certainly offer exciting possibilities for the future. A fourth observation is that we now have genuine process models of the way structured knowledge, that is so vital to cognitive development, is acquired. Fifth, a good deal of common ground has emerged from neo-Piagetian models about what constitutes conceptual complexity in cognition. This yielded a complexity metric, based on the number of related variables in a cognitive representation, that enables tasks to be compared for complexity, and tasks of equivalent complexity to be recognized, independent of domain or methodology. Furthermore it can be applied, not only to cognitive development, but to cognition generally. This is an illustration of the way cognitive development research can contribute to general cognition and cognitive science.

It is always nice to finish with a single principle that captures the essence of what has been learned from an extended body of research. One such principle emerges jointly from cognitive and neuroscience research. This is that with maturation and development, there is an increase in the ability to process complex information. Research that we have reviewed indicates that the anterior prefrontal cortex is specialized for coordination and integration of information, and for processing complex relationships. The anterior prefrontal cortex is also late developing, so we would expect that these functions will improve due to maturation. Basic cognitive functions, including associative learning and many perceptual functions, will be as efficient in infants and young children as in older children and adults. But complex processing, such as occurs in reasoning, as well as in language production and comprehension, is a factor that will become more efficient with age. This improvement will be influenced by both maturation and experience, because experience also contributes to neurocognitive development. However it means that simple conceptions that cognitive development depends on maturation or on learning, are imprecise and misleading. Some functions mature early, whereas complex cognitive functions mature later, and this needs to be taken into account in educational and child-rearing contexts.

References


