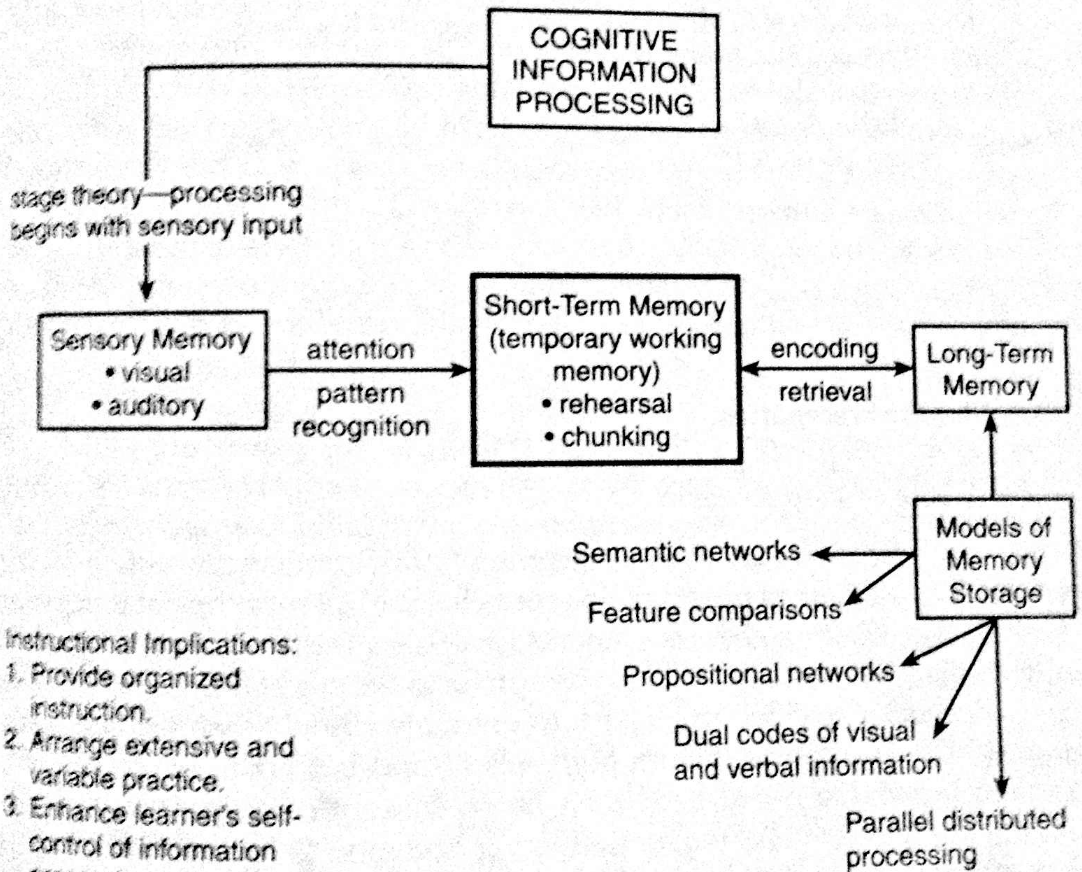
Part III: Learning and Cognition

3

Cognitive Information Processing



Instructional Implications:

1. Provide organized instruction.
2. Arrange extensive and variable practice.
3. Enhance learner's self-control of information processing.

Overview of the Information-Processing System
The Stages of Information Processing
The Flow of Information During Learning
Sensory Memory
Selective Attention
Automaticity
Pattern Recognition and Perception
Working Memory
Rehearsal
Encoding
Long-Term Memory
Representation and Storage of Information
Network Models of LTM
Feature Comparison Models of LTM
Propositional Models of LTM
Parallel Distributed Processing (PDP) Models of LTM
Dual-Code Models of LTM
Retrieval of Learned Information
Recall
Recognition
Encoding Specificity
Forgetting
Implications of CIP for Instruction
Providing Organized Instruction
Arranging Extensive and Variable Practice
Enhancing Learners' Self-Control of Information Processing
Conclusion
A CIP Look at "Kermit and the Keyboard"
Theory Matrix
Suggested Readings
Reflective Questions and Activities

Consider these scenarios.

• *A Tale of Two Readers*

Sarah lives in a small rural community and participates nightly in the county's adult literacy program. She reads haltingly, sounding out unfamiliar words. The selection she has chosen to work on this week is a simple tale about village life, and she is able to comprehend the gist of the text quite easily.

Rosemary decided to go back to graduate school when the last of her three children graduated from high school and left home for college. Although her children had used their home desktop computer regularly for school assignments, Rosemary had never bothered to learn. Now, some of her courses required access to the Internet, so she was forced to purchase a modem. The salesperson (and her classmates) assured her that hooking it up and using it was a simple matter. Unfortunately, an operating problem sent Rosemary to the manual about the modem, where she attempted to make sense of sentences such as, "The primary application for the local digital loopback is to permit a modem that is not CCITT V.54 compatible to engage in a remote digital loopback test with your modem."

• *The Mechanic and the Web Surfer*

Marcy arrived to pick up her car, which had been in for service, and her mechanic, Wes, explained the repairs that had been made in addition to the routine oil service. One of the reasons Marcy liked this particular shop was that Wes never talked down to her but took the time to explain what was wrong with the car when it needed to be fixed. In this case, Wes said, the steering damper and center link had to be adjusted, and the noise she had reported was coming from worn bushings around the tie rod ends in the suspension. Marcy nodded in understanding. As she prepared to leave, she and Wes chatted about an incident reported in the paper concerning a hacker who had shut down the local Freenet. Wes mentioned that his wife enjoyed using her account to e-mail friends and relatives all around the United States. He, on the other hand, didn't quite understand how the Internet worked and had become concerned after the hacker incident. Marcy, who enjoyed Web surfing herself, stayed a few moments longer to give Wes a basic lesson on the Internet.

Arriving at their respective homes that evening, Marcy and Wes had remarkably similar conversations with their spouses. In response to his question about her car, Marcy told her husband, "Oh, they fixed something on the steering, and that squeak is being caused by some rod rubbing against something or other. Nothing to worry about." Her husband shook his head; why did he even ask? To his wife, Wes said, "One of my customers today told me all about computers and e-mail and that stuff." "What about it?" his wife wanted to know, but unfortunately, Wes couldn't remember anything more specific.

Before proceeding further, reflect momentarily on the behaviorist perspective discussed in the previous chapter. How might a behaviorist account for the behaviors exhibited in these two scenarios? How is a complex behavior such as reading acquired? Why did Marcy and Wes experience such difficulty in recalling to their spouses what they had been told earlier in the day? Questions similar to these pose problems for behaviorism. And although behaviorism had dominated American psychology for half a century, it was to be supplanted by cognitive challenges.

Remember that the study of cognition was not new to psychology. Before radical behaviorism had gained such a stronghold on psychological research and theory, Tolman used cognitive maps to explain purposive behavior in rats, and Hull relied on a number of cognitive mediators between stimulus and response. Pavlov, as well, had introduced the concept of the "second signal system" to account for language learning. Vygotsky had launched his theory of how inner speech functions as a cognitive mediator explicitly in reaction to American behaviorism. Moreover, Gestalt psychologists in Germany had proposed that organizational processes in cognition were important to perception, learning, and problem solving. What was new in American psychology was the computer metaphor adopted for conceptualizing cognition.

The birth of computers after World War II provided a concrete way of thinking about learning and a consistent framework for interpreting early work on memory, perception, and learning. Stimuli became inputs; behavior became outputs. And what happened in between was conceived of as information processing. Today, what is known as cognitive information processing (CIP) is in reality an integration of views developed from a variety of perspectives.

> Like the traditional cognitive view, the CIP model portrays the mind possessing a structure consisting of components for processing (storing, retrieving, transforming, using) information and procedures for using the components. Like the behavioral view, the CIP model holds that learning consists partially of the formation of associations. (Andre & Phye, 1986, p. 3)

Overview of the Information-Processing System

According to the cognitive information processing view, the human learner is conceived to be a processor of information in much the same way a computer is. When learning occurs, information is input from the environment, processed and stored in memory, and output in the form of some learned capability. Adherents of the CIP model, like behaviorists, seek to explain how the environment modifies human behavior. But unlike behaviorists, they assume an intervening variable between environment and behavior. That variable is the information processing system of the learner.

Most models of information processing can be traced to Atkinson and Shiffrin (1968), who proposed a multistore, multistage theory of memory. That is, from the time information is received by the processing system, it undergoes a series of transformations until it can be permanently stored in memory. This flow of information, as it is generally conceived, is shown in Figure 3.1. Displayed in the figure are the three basic stages of the proposed memory system—sensory memory, short-term memory, and long-term memory—along with the processes assumed to be responsible for transferring information from one stage to the next. Let us briefly consider what these stages are and how they are believed to function.

The Stages of Information Processing

Sensory memory represents the first stage of information processing. Associated with the senses (vision, hearing, etc.), it functions to hold information in memory very briefly, just long enough for the information to be processed further. For example, imagine yourself in a dark, unfamiliar room. You strike a

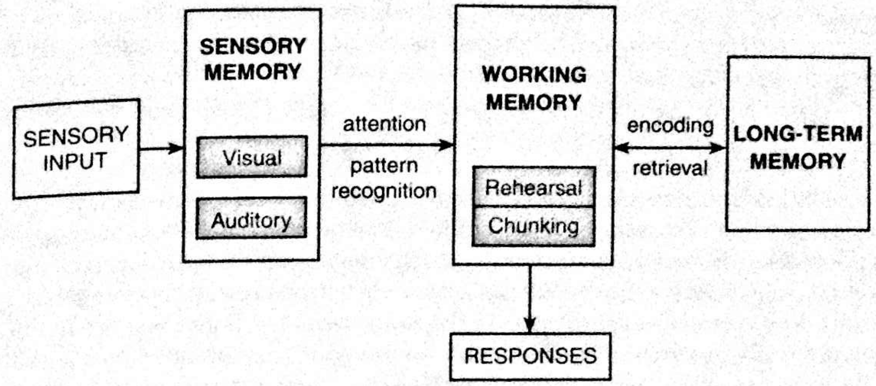


FIGURE 3.1 ***The Flow of Information as Generally Conceptualized in Information-Processing Theory***

match, which flares briefly and then goes out. In the split second after the match has gone out, you retain a visual after-image of the room, which stays with you just long enough for you to determine where the door or light switch is located. There is a separate sensory memory corresponding with each of the five senses, but all are assumed to operate in essentially the same way.

Working memory, also called short-term memory or short-term store (Atkinson & Shiffrin, 1968, 1971), is the stage at which further processing is carried out to make information ready for long-term storage or a response. Originally a unitary construct, working memory is generally thought to have independent processors for individual sensory modes (Baddeley, 1992). Working memory has been likened to consciousness. When you are actively thinking about ideas and are therefore conscious of them, they are in working memory.

Working memory not only holds information for a limited amount of time, but also holds a limited amount of information. In other words, you can think about only a few ideas at one time or read and understand relatively few phrases at once. With very long and complex sentences, for example, the reader has typically forgotten the beginning of the sentence by the time the end of it is reached. You can well imagine the effect on comprehension and recall that this limited capacity for keeping things in mind will have.

The **long-term memory** represents a permanent storehouse of information. Anything that is to be remembered for a long time must be transferred from short-term to long-term memory. Although forgetting is a phenomenon we have all experienced (and will be discussed in detail later in the chapter), it is assumed that once information has been processed into long-term memory, it is never truly lost. As for capacity, despite the protests of many children, long-term memory cannot be filled up. As far as we know, long-term memory is capable of retaining an unlimited amount and variety of information.

Table 3.1 displays a summary of the three stages of information processing that may help you keep their properties in mind as you progress through this chapter.

The Flow of Information During Learning

As indicated earlier, information is transformed—or processed—as it passes from one stage of memory to the next. What are the processes assumed to be responsible for these transformations? Let's examine a particular example from A Tale of Two Readers to trace what may happen during learning. Suppose Sarah comes to this sentence in the story she is reading: "Visitors to the town are always struck by the beauty of its wide, azalea-lined avenues." The letters on the page stimulate Sarah's visual sensory register, which receives and briefly records a representation of the information as it originally occurred. Then, familiar shapes of letters and words are perceived as pattern recognition takes place. It is at this point that the process of attention also exerts an effect. An unfamiliar word may cause processing to slow, because added attention must be paid to individual letters rather than whole words.

Upon entering working memory, the information is coded conceptually, i.e., takes on meaning. Meanings of the individual words are retrieved from long-term memory to assist Sarah in constructing a representation of the whole sentence. Since the sentence is more than a few words, internal rehearsal may also occur to preserve the first few words in memory while the end of the sentence is being perceived.

Finally, in order for the information to be processed into long-term memory, Sarah must encode its meaning. This means that the representation

TABLE 3.1 *Summary of Memory Stages*

	Stages		
Properties	*Sensory Register*	*Short-Term Store*	*Long-Term Store*
Capacity	Large	Small	Large
Code	Literal copy of physical stimulus	Dual code —verbal —visual	Episodic/semantic
Permanence	0.5 seconds	20–30 seconds	Permanent
Source	Environment	Environment and prior knowledge	Effective encodings from STS
Loss	Decay	Displacement or decay	Irretrievability

Source: From McCown, R. R., Driscoll, M. P., & Roop, P., *Educational psychology: A learning-centered approach to classroom practice.* Boston: Allyn & Bacon, 1996. Reprinted with permission.

she constructs of the sentence must be meaningful and make connections with related knowledge already in long-term memory. For example, her previous experience with azaleas and wide streets may allow her to construct an image of what this sentence describes. Her image then becomes a useful retrieval cue when she is asked to recall what she has read.

It may be evident from this example that processing does not truly occur in the unidirectional, linear way in which it is often depicted (e.g., in Figure 3.1). Instead, the representation Sarah constructs of the sentence is determined both by the information itself (data-driven, bottom-up processing) and by her prior knowledge (conceptually driven, top-down processing). The degree to which either type of processing dominates seems to depend on the nature of the learning task itself and the amount of prior knowledge the learner brings to it.

Little has yet been said about the control processes influencing information flow. Whether these are thought of as comprising a system component (e.g., Gagné, 1985; Andre & Phye, 1986) or as processes modifying information flow within and between components (e.g., Atkinson & Shiffrin, 1971), they have the same effect. In some way, an executive monitor keeps track of the information flow and makes decisions about processing priorities. This may occur in a conscious, strategic fashion or in an unconscious, automatic way. For example, Sarah may have very deliberately chosen to associate an image with the sentence she read, because she has found imagery to be a very effective study strategy. On the other hand, suppose the story had previously described only camellias adorning villagers' lawns. Sarah may then have developed an expectation that could cause her to mistakenly perceive "camellias" instead of "azaleas" in the target sentence. In either case, a control process has modified the information flow and what was ultimately understood and learned.

The sections that follow focus on each of the major stages and processes of the human processing system. As you read them, keep in mind two things. First, the computer provided a concrete metaphor for human information processing and, thus, a language for describing and integrating a variety of learning phenomena. Second, for instruction to be meaningful and relevant, it must build upon learners' prior knowledge and help learners to construct cognitive connections between what they already know and what they are being asked to learn.

Sensory Memory

The existence of some sort of perceptual store in the information-processing system that registers information and holds it very briefly was demonstrated in a series of experiments conducted by Sperling (1960). Sperling presented subjects with a visual array of twelve letters (arranged in three rows of four letters each), such as the one shown in Figure 3.2. He flashed the array on a screen for 50 milliseconds (one-twentieth of a second) and then asked subjects to report what letters they had seen. Even with such a brief exposure, subjects could consistently report three or four letters accurately.

Although this result seemed to indicate a limited processing capacity, Sperling was able to show that, in fact, all of the letters had entered sensory memory. He did this by using a partial report technique. That is, he used a high, medium, or low tone to signal to subjects which row of the array they were to report. Instead of a relatively poor performance (three or four of twelve letters), subjects showed remarkably good performance, reliably reporting three or four letters in the row (so, three or four of four) no matter which row was signaled. It appears, then, that sensory memory is temporally, rather than visually, limited. In other words, a great deal of visual information registers, but it decays very rapidly without further processing, within a quarter of a second, according to Sperling's experiments.

Relatively little is known about the sensory memories corresponding to the other senses, but they are presumed to function in a similar way. Darwin, Turvey, and Crowder (1972) replicated Sperling's results with the auditory system. They found, however, that the auditory sensory memory (or echo) lasted longer than the visual sensory memory (or icon), typically up to 4 seconds under partial report conditions. An explanation for this difference is thought to lie in the requirements for speech perception. In other words, sounds must remain in sensory memory long enough for them to be combined with other sounds so that speech may be understood.

Sperling's use of the partial report technique also illustrates the effect that attention has on information processing. The tone served as a cue to focus attention on a particular part of the display so that it could be processed further and recalled. Attention is a process that has been conceptualized in a variety of ways. Instructors admonish students to pay attention in class, but they also adopt measures to focus students' attention on particular features of instruction. Either way, a student who is not attentive misses some of the information to be learned.

Cognitive psychologists, noting that some information always seems to be lost in processing, initially thought that attention acted as a bottleneck or

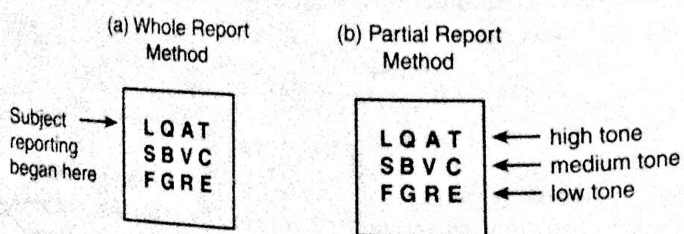


FIGURE 3.2 *Visual Displays Similar to Those Used by Sperling (1960)*

filter preventing information from entering the system (e.g., Broadbent, 1957). Treisman (1960) showed, however, that attention is not an all-or-none proposition and suggested that it serves to attenuate, or tune out, stimulation. Her ideas are easily illustrated by thinking about what happens at parties. You may be attending to one conversation, ostensibly unaware of what else is being said around you. But when you hear your name spoken or someone else talking about a topic that interests you, your attention shifts. Apparently enough information was being processed to prompt you to react and pay closer attention to the secondary source.

Researchers have come to view attention as a resource with limited capacity to be allocated and shared among competing goals (cf. Kahneman, 1973; Grabe, 1986). This suggests that learners have some control over the process and may selectively focus attention to meet certain ends. It is also true, however, that some tasks require relatively little attention and may be accomplished effortlessly and automatically. The concepts of selectivity and automaticity are important aspects of attention that hold implications for instruction. Let us consider each in turn.

Selective Attention

Selective attention refers to the learner's *ability to select and process certain information while simultaneously ignoring other information.* The extent to which individuals can spread their attention across two or more tasks (or sources of information) or focus on selected information within a single task depends upon a number of factors. The most obvious, perhaps, is the meaning that the task or information holds for an individual. Your name spoken in a crowded room catches your attention because it is highly meaningful to you.

Second, similarity between competing tasks or sources of information makes a difference. It is hard to listen to two conversations at the same time when both speakers are the same sex and are speaking in a similar tone and volume. Imagine the poor student, for example, who is trying to listen to the teacher at the same time a classmate talks in her ear. Similarly, a learner may enjoy studying to classical music but find her concentration slipping when vocal music is played.

Task complexity or difficulty is a third factor that influences attention. Simple tasks, such as winding yarn into a ball, require relatively little attention and are easily done at the same time as other things. Watching a lighthearted TV comedy, putting together a jigsaw puzzle, and talking to your family about tomorrow's schedule are probably all tasks that can be accomplished simultaneously. But reading a medical history for purposes of diagnosis or assembling an intricate set of electrical circuits demands more complete and focused attention. A task may also demand more attention when it is something about which the learner has little prior knowledge. For example, a post-baccalaureate student taking his first course in learning theory is likely to find it necessary to pay close attention to both his instructor and the textbook.

Finally, the ability to control attention, in both a general and specific sense, appears to differ with age, hyperactivity, intelligence, and learning disabilities (Grabe, 1986, p. 66). For example, attention deficit disorder is a condition afflicting a small proportion of preadolescent children. They seem to be unable to focus attention or to turn off irrelevant stimulation. As a result, their school performance typically suffers.

How, then, is attention best managed in instructional situations? To influence attention or alertness of students during the course of a classroom lesson, Good and Brophy (1984) recommended that instructors employ standard signals (e.g., "Let's begin," "Back on task!," turning the lights on or off). A third grade teacher of my acquaintance uses a maraca to gain the attention of all students when they are working in pairs or small groups. Because he has used that signal from the first day of class, students know when they hear it that they are to stop whatever they are doing and look at him for direction.

When it is important to focus students' attention on certain aspects of the instructional materials, stimulus features can be highlighted through the use of color or type of print (in textual materials), voice inflections or gestures (in classroom presentations), and novelty. To emphasize the different sorts of roles that computer consultants often play, for example, a community college teacher wears different hats during his lecture, each one representing a different role.

Finally, Grabe (1986) reviewed ways in which learners themselves may be taught to stay on task and selectively attend to important features of instruction. He indicated that two things appear to be important: (1) Learners should be taught to take more time in responding to a learning task (to reduce impulsiveness), and (2) they should be given a strategy for focusing attention and allowed to practice that strategy (p. 74). Certain games that require attention, e.g., *Concentration* or *Simon Says*, can be used to help students develop better attending skills.

Automaticity

When tasks are overlearned or sources of information become habitual, to the extent that their attention requirements are minimal, **automaticity** has occurred. Driving a car provides a good example of the distinction Shiffrin and Schneider (1977) made between automatic and controlled processing. For the most part, the driving task is automatic, enabling the driver to listen attentively to a radio program, for example. But when traffic is heavy or something unusual occurs to demand the driver's attention, driving shifts to a controlled process. The driver then must pay much closer attention to driving and fails to hear what is being said on the radio.

LaBerge and Samuels (1974) have developed a theory to account for automatic processing in reading. They believe decoding words should be so automatic for readers that they can concentrate their attention on comprehending the meaning of what is read. In A Tale of Two Readers, for example, Sarah has learned to decode but has not yet learned the skill to the point where it is automatic. As a result, her reading is slow and fraught with difficulty. Rosemary, on the other hand, may decode automatically most of the time, but here faces unfamiliar information that makes her comprehension of the meaning difficult. As a result, she, like the driver in traffic, must shift from automatic to controlled processing in order to decode the unfamiliar words.

To develop automatic decoding skills in readers, researchers have explored a number of possibilities, including extended word identification practice as part of the regular text-reading curriculum (e.g., Beck, 1981, 1983). More recently, researchers have become encouraged by the potential of the computer for providing many different types of word tasks in an engaging environment (Perfetti & Curtis, 1986). It may also be useful for teachers to include read-aloud activities (such as reading and answering questions) after learners have read silently. Readers' sensitivity to different kinds of scripts can impair their comprehension, but such impairment seems to be obviated by reading aloud during rereading (Jacoby, Levy, & Steinbach, 1992).

Once reading is automatic, however, what readers will comprehend and remember from text depends upon how they allocate their attention as they read. Readers will generally allocate greater attention to important elements in a text (Anderson, 1982). They determine importance based on the purpose for which they are reading as well as features of the text that signal something is important.

As noted in the previous section, the reader's attention can be directed by typographical cues in the text (e.g., boldface print, capitalization [Glynn & Divesta, 1979]), as well as the presence of titles (Kozminsky, 1977), specific phrases (e.g., "an important cause of..." [Armbruster, 1986]), and idea unit structure (Kintsch & van Dijk, 1978). Idea unit structure refers to the placement of main ideas and supporting details within a paragraph. Ideas that appear high in the structure are more likely to be attended to and remembered than details buried deep within a paragraph. Writers of instructional texts, then, are well advised to employ these features to direct learner attention to the important, to-be-learned information.

Readers, on their own, also differentially allocate attention according to the purpose for which they are reading. Reading a novel, for example, typically involves reading for the gist of a story, and readers may be hard pressed to recount very specific details when they are finished. Reading a textbook or technical manual, on the other hand, is done with a specific purpose in mind—to locate and learn important information. Assigning instructional objectives (Klauer, 1984) or inserting questions in the text (Andre, 1979) has proven effective for helping students focus their attention on specific text information.

Automatization of other basic skills besides reading (such as the rules of arithmetic operations and grammar) is considered to be a desirable educational goal for the primary grades (Gagné, 1983; Bloom, 1986). By extension, one can also see the usefulness of learning certain tasks as adults to automaticity. Pilots must react automatically to a variety of information sources in the cockpit. Astronomers automatically process patterns of stars in the search for anomalies that might be signs of new stars or other astral phenomena. And detection of signs of abuse is probably automatic for skilled therapists.

Pattern Recognition and Perception

Just attending to information is not enough to ensure its further processing, however. One might say that attention is necessary but not sufficient; information must also be analyzed and familiar patterns identified to provide a basis for further processing. **Pattern recognition** refers to the process whereby *environmental stimuli are recognized as exemplars of concepts and principles already in memory.* This recognition is preconceptual, something like finding a shape that matches a stencil without identifying what the shape or stencil pattern actually represents.

Pattern recognition is a particularly difficult process to model in the human information-processing system, and, consequently, several different models have been proposed. Each carries particular implications for how the process operates and for what form information is represented in memory. Briefly, template matching assumes that mental copies of environmental stimuli, or templates, are stored in memory. Pattern recognition, then, consists of simply matching the incoming information to the appropriate template in memory. Although this seems intuitively appealing, look at Figure 3.3 and consider what this means for a template-matching model of pattern recognition. In order for you to recognize all of those figures as representations of the letter *A*, you would have to have templates in memory to match each one. For obvious reasons, this model fails as a reasonable account of human pattern recognition.

According to an alternative, prototype model, what is stored in memory is not an exact copy of a stimulus, but rather an abstracted, general

A a A **A** **a** a A a

FIGURE 3.3 *Variations of the Letter A*

prototype. Pattern recognition in this case involves comparing the incoming information with the prototype. If a close enough match is found, then the incoming stimulus is recognized as an example of the class of objects or events represented by the prototype. Thus, all of the letters in Figure 3.3, for example, are similar enough to the assumed prototype to be recognized as *A*s.

The prototype model has become popular for explaining pattern recognition, primarily because of evidence that suggests we tend to store prototypic concepts in memory (see, in particular, Eleanor Rosch's work [1973, 1975]). For example, asked to indicate what color comes to mind in response to the verbal stimulus *red,* you are likely to choose a color that is close to fire-engine red. Similarly, reading about Olympic athletes or shore birds tends to evoke general ideas about these concepts rather than specific, previously encountered examples.

A third model of pattern recognition, called feature analysis, presumes that specific, distinctive features are stored in memory. Incoming information is then analyzed for the presence of these features. To consider the letter *A* one more time, its defining features might include the two sides, the angle at which they are joined, and the horizontal connecting line. All stimulus letters would be analyzed for these features and, if found, would be identified as *A*s.

Feature analysis, like the prototype model, is supported by experimental evidence and together, the two models have influenced pedagogical recommendations for concept learning. Tennyson and Cocchiarella (1986) proposed a model for teaching concepts that calls for presenting, first, a best or prototypic concept example followed by a variety of examples that differ from the prototype in systematic ways. The examples help learners to abstract the meaningful dimensions of the concept and determine which features are critical and invariant and which are nonessential and variable across examples.

To see how this model might work, consider one of the concepts from the previous chapter: positive reinforcement. A best or prototypic example might be one in which positive reinforcement is shown with animals and the use of a primary reinforcer. Then, additional examples could be explored in which positive reinforcement is demonstrated with children in school or adults at work and the use of secondary or social reinforcers. Or consider how a medical student might learn to distinguish diseased from normal cells. With stained slides showing clear examples of each for comparison, the student could examine other slides bearing cells that show a range of what is considered normal characteristics and a range of what is identified as diseased.

Although the feature comparison and prototype models account well for most instances of pattern recognition, they also are unable to account for why certain patterns are recognized even though all the features are not

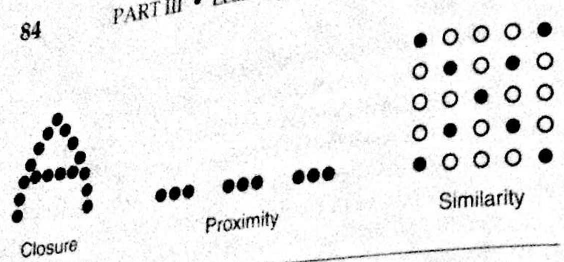


FIGURE 3.4 *Gestalt Principles of Organization*

present or they fail to resemble their prototype. For example, a degraded copy of the letter *A*, as might be seen on a badly eroded tombstone or a poorly produced overhead transparency, is still recognized as an *A*.

To explain this and other perceptual phenomena not easily handled with the prototype or feature analysis models, we rely on principles of organization, context, and past experience. Gestalt psychologists, in studies dating from the early twentieth century, demonstrated that human perception tends to involve "going beyond the information given" in order to construct a meaningful interpretation. That is, the way in which stimuli are organized will prompt the viewer to see them in certain ways, apart from what is actually there. For example, look at the pictures displayed in Figure 3.4. What do you see? Chances are, you did not say, "Just a bunch of dots." The principle of closure prompts us to close up the spaces between the dots in Figure 3.4 (left) and to perceive the figure as an "*A*". Due to proximity, we view the dots in Figure 3.4 (center), not as *nine dots*, but as *three sets of three dots*. Finally, similarity dictates that similar units will be perceived as one, so that we do not see black and white dots in Figure 3.4 (right), we see a black *X*.

The effect of context on pattern recognition can be illustrated by reference to the tombstone and overhead transparency mentioned earlier. In those instances, why is it likely for the degraded letter to be perceived as an *A*? Presumably, the reason is that clues to its identity exist in the context that surrounds it. Other, more easily perceived letters suggest what words are on the tombstone and transparency. Once the word containing the degraded letter has been determined, the identity of the letter is obvious. Figure 3.5 illustrates how context is used to resolve some perceptual ambiguity. The figure in the center could be either the letter *B* or the number *13*. Which will be perceived depends on whether the other letters in the row or the figures in the column are used to provide contextual clues.

Past experience, or prior learning, is the last factor to be considered for its effect on pattern recognition. This refers to the simple fact that what has been learned or experienced previously will have some impact on what is perceived in later situations. A good illustration of this can be seen in the

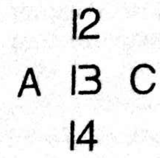


FIGURE 3.5 *Context Effects on Perception*

Stroop effect. An individual is shown a series of color words (e.g., blue, green, or red) that are printed in different colors and is asked to "name the colors as quickly as you can." What happens is that the person has great difficul ty in identifying the colors of the words, tending instead to read the words themselves. Knowledge of color words, coupled with reading skill, interferes with one's ability to perceive the colors. The same would hold true for proofreading; one has a tendency to read the words as they should be typed rather than as they actually are.

Solving problems can also require overcoming the effects of past experience on perception. In other words, some problem situations must be perceived in a new way in order for a solution to be reached. In Kohler's (1925) experiments with a chimpanzee, for example, bananas were placed just out of the chimpanzee's reach with a stick near at hand. In order to get at the bananas, the chimpanzee had to perceive the properties of the stick as affording its use as a tool to knock the bananas within reach. Similarly, solving an insight problem such as "If the lily pads on a pond double every day, and the pond is completely covered on the 100th day, on which day is it half-covered?" requires thinking of the problem in terms of logic rather than math.

Although little is known about how people come to be proficient at casting problems in a new light in order to solve them, there is evidence to suggest that practice on many different kinds of problems may help (Sternberg & Davidson, 1983). Practice with a variety of problems can make learners more aware of the role of context in problem solution and thus more open to the consideration of alternate assumptions.

The influences of past experience and context on perception can also come together in expectations about students. It has been well documented that teachers' expectations of students may affect their evaluations of student achievement, as well as their own behavior toward students (e.g., Good, 1987). In other words, expecting a student to be a problem in class can predispose the instructor to perceiving more problem behaviors. Similarly, a student with a reputation for high achievement is more likely to be perceived in that light.

The expectations themselves may develop from previous experiences of the teacher, from the immediate context, or both. For example, the teacher has learned to associate, and therefore comes to expect, certain behaviors with high- and low-achieving students, males and females, or well- and poorly behaved children. But context also plays a part. Teachers may expect less of the same individual in a generally high-achieving class than in a class that performs less well overall.

Although the self-fulfilling prophecy (Rosenthal and Jacobson, 1968) has had a considerable influence in schools over the past 20 years, recent evidence has shown that what teachers do (or fail to do) matters more than what teachers expect with regard to student achievement. Goldenberg (1992) described two cases of paradoxical expectancy in which the children's first grade, year-end achievements were in marked contrast to what the teacher had expected. He concluded in one case that "The teacher had failed to take corrective action when she should have *because she had expected* [the student] to do well on her own" (Goldenberg, 1992, p. 539). In the other case, "*in spite of the teacher's low expectations* for [the student's] success, the teacher took actions that appear to have influenced [her] eventual first-grade reading achievement.... Low expectations were clearly evident, but they were irrelevant in determining the teacher's actions" (p. 539). Although expectations can have an influence on teacher behavior, then, they do not always matter. What appears to be more important is whether the instructor monitors student achievements and takes corrective action as necessary.

Sensory memory, attention, and pattern recognition, while important, obviously tell only part of the story. When learners have paid sufficient attention and pattern recognition of selected portions of the stimulus has occurred, a great deal more processing is still required for the information to become a meaningful and permanent part of memory. The next stage of activity occurs in working memory.

Working Memory

Information selected for further processing comes to the working memory. At this stage, concepts from long-term memory will be activated for use in making sense of the incoming information. But, as indicated earlier in the chapter, there are limits to how much information can be held in working memory at one time and for how long information may be retained there, unless, of course, something is done to increase capacity or duration in some way.

In a now classic study of short-term memory, George Miller (1956) demonstrated that about 7 ± 2 numbers could be recalled in a digit-span test. This test consisted of reading subjects a list of numbers and asking them to immediately repeat what they had heard. With seven items being the typical memory span, is it any surprise that local phone numbers are exactly seven digits? Miller also whimsically wondered whether there are magical qualities to the number 7; after all, there are "the seven wonders of the world, the seven seas, the seven deadly sins, the seven daughters of Atlas in the Pleiades, the seven ages of man, the seven levels of hell, the seven primary colors,

the seven notes of the musical scale, and the seven days of the week" (Miller, 1967, pp. 42–43).

Despite Miller's whimsy, seven bits of information have been shown to constitute the memory span for a great variety of materials. Moreover, each bit of information can vary tremendously in size. A ten-letter word, for example, may be one bit, along with a six-word sentence. Discovery of this fact has led to the notion that *working memory capacity may be increased through creating larger bits*, known as the process of **chunking.** Take, for example, the span of letters shown below.

JFKFBIAIDSNASAMIT

As individual letters, they more than exceed working memory capacity. But as five chunks—JFK, FBI, AIDS, NASA, and MIT—they are easily processed.

What this has been taken to mean for instruction is that learning tasks should be organized so that they can be easily chunked by the learner. This may be as simple as breaking complex tasks into manageable steps, as in a science experiment, or presenting discrete bits of information to be studied and practiced, as in the frames of a computer-based tutorial lesson. In addition, issues in political science that involve very complex arguments, for example, will be better understood if the arguments are broken down and examined bit by bit.

How chunks of information are actually stored in working memory has been likened to a series of slots, with each chunk taking up one slot. As new chunks come into memory, they push out those that were previously occupying the available spaces. This is now the accepted explanation for the serial position effect known as recency. In the serial position task, subjects are given a list of words or nonsense syllables to learn. Typically, fifteen or twenty items are presented at a rapid rate, and immediately following the last item, subjects recall as many as they can. You can guess which ones they recall best—the items at the end of the list or those seen most recently. It was assumed, then, that later items on the list pushed out of memory those that had been seen first. There was simply not enough room for them all.

To determine the duration of working memory, Brown (1958) and Peterson and Peterson (1959) presented subjects with sets of three letters they were to recall after brief intervals. What seems like an easy task becomes much more difficult when rehearsal is prevented during the retention interval. That is, subjects had to count backwards by threes from a given number until the retention interval was up. Results indicated that memory for the letters was still good after only 3 seconds, but after 18 seconds, decay was nearly complete. Given individual differences, it is generally accepted that unrehearsed information will be lost from working memory in about 15 to 30 seconds. In the days of the rotary dial, this is about the same amount of time it took to dial a number and get a busy signal!

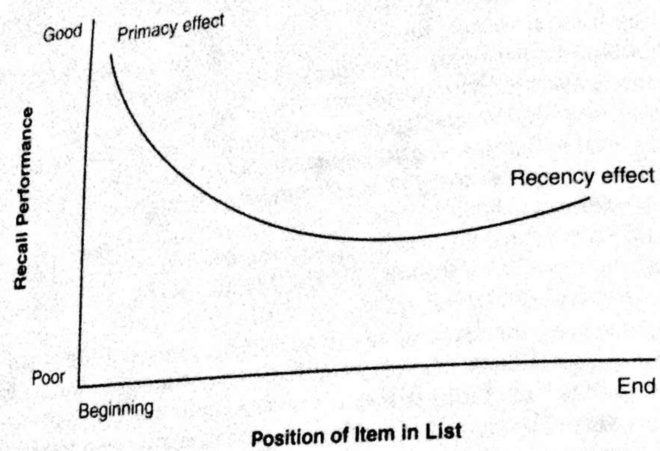


FIGURE 3.6 *Serial Position Curve*

In order to prevent the loss of information from working memory, and to ensure its being transferred to long-term storage, two processes are necessary: rehearsal and encoding.

Rehearsal

When you repeat a phone number to yourself over and over while waiting to use the phone, you are engaged in rehearsal. Some would call this maintenance rehearsal because the repetition serves to maintain the information in the working memory for some designated period of time. Once you have made the call and reached your party, you no longer have the need to maintain the phone number in the short-term store.

Rehearsal has been used to explain the primacy effect of the serial position curve. When items are presented as described earlier, but at a slower rate, subjects remember not only the last items on the list, but the first ones as well (Figure 3.6). You can imagine why. With only a few items in memory at the beginning of the list, subjects have time between items to rehearse all the items they have seen. As more items crowd in, however, the rehearsal task becomes more difficult, so that items in the middle of the list receive less practice. As before, items at the end are recalled well because they are still in working memory at the time of recall.

Whereas recency and primacy effects are ostensibly associated with short-term memory, there are anecdotal data to suggest that something similar goes on even after information should be in long-term memory. For example, if a pop quiz is given after a typical 50-minute lecture, chances are students will remember best what was discussed in the first 10 minutes of class and in the last 10 minutes before the quiz. Likewise, most journalists adhere to the maxim that important information should go at the beginning and end of their articles, because these are the paragraphs best remembered by readers. These phenomena have led some researchers to question the dual-stage nature of memory and to propose instead some sort of intermediate memory or a continuum from short-term to long-term memory.

Finally, for information to reach a relatively permanent state in long-term memory, maintenance rehearsal is not enough. Learners will argue that simple repetition is an effective means for them to remember something for a long time. In the case of highly overlearned material, such as arithmetic facts, spelling words, or a memorized script, they are probably right. But repetition of more complex and meaningful information will not ensure its being fully processed into long-term memory. Elaborative rehearsal, or encoding, will.

Encoding

Encoding refers to the process of relating incoming information to concepts and ideas already in memory in such a way that the new material is more memorable. Left to their natural inclinations, humans will always try to make things meaningful, to fit some new experience into the fabric of what they already know. We have already seen the evidence of this in perception and attention. Encoding serves to make permanent what these processes have initiated.

Studies demonstrating the various ways in which encoding may take place are too numerous to review in any comprehensive fashion here. But it is useful to consider briefly the major types of encoding schemes that have been investigated. The concept of organization, to begin with, has long been of interest to psychologists and educators alike. Bousfield (1953) found that people will group related pieces of information into categories in order to learn and remember them better. Even when information is seemingly unrelated, learners will impose their own, subjective organization on the material in order to learn it (Tulving, 1962). To assist learners in organizing material meaningfully, outlines (Glynn & Divesta, 1977), hierarchies (Bower et al., 1969), and concept trees (see the examples provided in Chapter 2 and later in this chapter [Tessmer & Driscoll, 1986]) have all proven effective.

Mnemonics and mediation (Matlin, 1983) provide other effective means for encoding. Learning a list of unrelated words, for example, is facilitated by linking the words together in the form of a story (Bower and Clark, 1969). The story then serves as a mediator to make the words on the list, which are meaningless by themselves, more memorable. This can be a helpful strategy for young children to use while learning to read. By themselves, single words may not have much meaning at first. But when children write stories incorporating certain words, they often find it easier to read and recognize these words later. Similarly, mnemonics such as ROY G BIV or "My Very Earnest Mother Just Showed Us Nine Planets" serve to aid in the learning and recall of the colors in the spectrum and the planets in our solar system (see reviews of mnemonic strategies by Higbee [1979] and Bellezza [1981]).

Finally, imagery can be a very effective means of encoding information. Studies have shown that pictures suggesting visual images (Levin & Kaplan, 1972) or simply instructions to form images related to text material (Kulhavy & Swenson, 1975) are effective in facilitating learning. Some teachers now find that combining this method with story creation, as described, can be a very powerful means for facilitating not only learning but motivation (D. Cooper, personal communication, September, 1992). Children "publish" their stories by drawing illustrations to accompany them. In so doing, they strengthen their understanding of words in a very personal, meaningful way.

Before leaving this topic, it is perhaps wise to point out that nearly any method of elaborative encoding is better for learning than is mere repetition of information. But which approach is best depends upon the learners and the material to be learned. Moreover, learners who have developed idiosyncratic but effective encoding strategies will not necessarily benefit from some strategy imposed by the instruction. For this reason, there has been considerable interest in determining how learners may be taught to develop and use their own strategies effectively (cf. Pressley & Levin, 1983; Levin & Pressley, 1986; Segal, Chipman, & Glaser, 1985).

Learners may be encouraged to invent their own mnemonics, for example. Instructors in a driving-under-the-influence program who attended a workshop I presented invented the acronym VOMIT to remind themselves of the effects of drinking on the driving task. (I no longer recall what the individual letters stand for, but no doubt they do! This just illustrates how individually effective mnemonics can be; what works for one learner may not for another.)

Self-questioning has also been investigated as a means for learners to encode information they hear in lectures or read in printed instructional materials. Sometimes learners ask themselves questions to aid in comprehending material, such as, "How does the meaning of this concept differ from what was discussed on the previous page?" Other questions, which call for drawing inferences, should help learners to integrate new information with what they have already learned.

In reviewing research on self-questioning as an encoding strategy, Snowman (1986) pointed out that some learners must be taught how to frame good questions if the strategy is to be effective. Some teachers do this as early as the second grade by asking their students, "What could you ask yourself to be sure you understand ___?," and then providing feedback on the students' responses (S. Briggs, personal communication, October, 1992).

But Ormrod (1990) speculated that it might be just a matter of students asking fact-based, low-level questions because they have learned to expect such questions on class examinations. Perhaps requiring learners to demonstrate inferential thinking in class and on assessments will prompt them to generate more inferential self-questions at encoding.

It may seem, in this discussion of working memory, that some aspects of permanent memory have already been touched upon, and indeed they have. It is virtually impossible to divorce the processes of working memory from those of long-term memory completely, because they are intimately related. Encoding, for example, by virtue of its role in transforming information as it passes from working to long-term memory, could be as easily discussed under the framework of the latter as the former. Encoding will continue to play an important role as we now consider what happens to information when it reaches long-term memory.

Long-Term Memory

Do you remember what you had for dinner last night? Or what you did on your birthday last year? Perhaps you recall a visit to another country where the most memorable events were your donkey ride down a steep embankment, the shopkeeper who offered you ouzo at nine o'clock in the morning, and the hotel manager who kept repeating, "So sorry. No reservation." Now consider how these memories differ from your knowledge that Albany, not New York City, is the capital of New York and that reading a weather map will tell you whether to expect rain in the next few days. Although these are all examples of information you retain in long-term memory, they differ in whether they represent specific experiences unique to you or general knowledge of the world that is shared by others.

Tulving (1972) was the first to make the distinction between episodic and semantic memory. He conceived of these as two information processing systems, each selectively receiving information, retaining certain aspects of that information, and retrieving the information as required. Episodic memory is memory for specific events, as when you remember the circumstances surrounding how you learned to read a weather map. Semantic memory, on the other hand, refers to all the general information stored in memory that can be recalled independently of how it was learned. For example, perhaps you cannot remember how you learned to read a weather map, because the circumstances surrounding the event were not particularly memorable. But you do remember the skill.

Although the two systems are related, it is semantic memory that most concerns educators. Generally, what is supposed to be learned in school, or indeed in any instructional situation, is semantic in nature. Before 1972, Tulving argued, most memory research concerned episodic learning. Since then, however, researchers have focused primarily on semantic memory, devising theories for how semantic information is represented in memory, how it is retrieved for use, and how it is forgotten. These questions provide the basis for discussion in the next several sections.

Representation and Storage of Information

How information is represented and stored in semantic memory is a central issue in the study of long-term memory (LTM) and one that has concerned researchers for centuries. Consider the difficulty of the task. Questions must be answered such as, What is the nature of the knowledge unit that is stored in memory? How are relations among these units represented? How can we account for individual differences in memory? Is there only one kind of knowledge unit, or are visual images substantively different from verbal propositions? Try to keep these questions in mind as some of the proposed answers are presented.

Network Models of LTM. One way to conceive of long-term memory is to think of it as a sort of mental dictionary (Klatzky, 1980), but instead of words being represented alphabetically, concepts are represented according to their associations to one another. For example, if I say "black," what comes to mind? I expect you said "white," which is closely associated with *black* by virtue of being its opposite. Other kinds of associations are obviously possible. A *canary* is a kind of bird, while *has gills* is a property of fish.

Network models assume the existence of nodes in memory, which correspond to concepts, i.e., things and properties. These nodes are thought to be interconnected in a vast network structure that represents learned relationships among concepts (e.g., Collins and Quillian, 1969).

Network models have the advantage of representing individual differences among learners, because individual learning histories presumably lead to different memory networks. These models also enable predictions, which can be easily verified by the performance of learners on certain memory tasks. For example, look at the partial network shown in Figure 3.7. That memory might be structured this way can be ascertained by asking subjects to respond to sentences such as, "A bird has wings," or "A blue heron is a fish." Since the concept *bird* points to the property *has wings* (assuming this was a learned relationship), the subject should say the first sentence is true. In the case of the second sentence, however, *blue heron* and *fish* cannot be directly connected, because the search process can only proceed in the direction indicated by the arrow. Thus, this sentence must be false.

In a similar fashion, predictions can be made about the speed at which subjects should be able to verify sentences as true. For example, learners should be faster in recognizing the truth of "A blue heron has long legs" than "A blue heron is an animal." In the first case, search had to proceed across

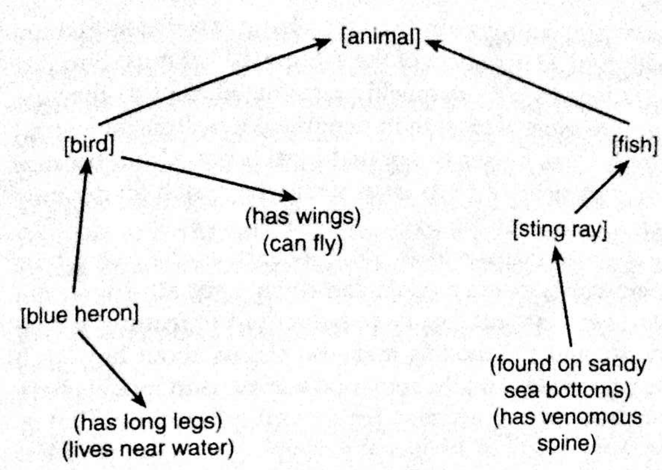


FIGURE 3.7 ***A Partial Network Representing Concepts Associated with Animal, in the Tradition of Collins and Quillian (1969)***

only one pointer; in the second case, two pointers, or levels of memory, are searched.

Predictions such as these were, in fact, confirmed by Collins and Quillian (1969), providing experimental support for the network models. But they also encountered some troubling findings. Subjects more quickly recognized a canary as a bird, for example, than a penguin as a bird, yet recognition times should be equal since the distance in both cases is the same. Typicality of concepts, then, presented a real difficulty for network models, which was to be overcome by feature comparison models of long-term memory.

Feature Comparison Models of LTM. Smith, Shoben, and Rips (1974) proposed that concepts in memory were not stored in interconnected hierarchies, as suggested by network models, but with sets of defining features. Association to other concepts is then accomplished through a comparison of overlapping features, hence, the label feature comparison models. For example, to verify "A blue heron is a bird," an individual would search all the characteristics of *blue heron* and all those of *bird*, and finding a sufficient overlap, would say the sentence is true.

Feature comparison models nicely explained the typicality effects so troubling to network models. Some concepts simply do not have clearly defined members; they are "fuzzy sets" in which some members are better examples of the concept than others. Thus, feature comparison models distinguished between defining and characteristic features. Defining features are those that a bird, for example, must have in order for it to be classified in that category. Characteristic features, on the other hand, are those that are usually associated with typical members of the category. That most birds fly is an example. Thus, canaries are more quickly recognized as birds than are penguins because they are more typical than penguins, which swim instead of fly. In a similar way, it takes longer to say that a bat is not a bird, because bats share features characteristic of birds even while the match on defining characteristics is poor.

Since there are a great many real world concepts of the fuzzy type (Kintsch, 1974), feature comparison models can seem very attractive. But they are not particularly economical, i.e., large collections of features would be required for learning, and the models make no claims about how such collections would be organized. Finally, semantic comparison models have been criticized for their failure to account for semantic flexibility. That is, context can cause certain aspects of a concept's meaning to be more or less prominent. If you hear, "Help me move the piano," you will probably think of it as a heavy piece of furniture, but the sentence, "You play the piano beautifully" emphasizes its musical aspect (Barclay et al., 1974).

Propositional Models of LTM. How different from one another are network and feature comparison models? In posing this question, Klatzky (1980) cited evidence that feature comparison models may in fact be rewritten as enhanced network models. Perhaps for this reason, the network has remained the primary metaphor for long-term memory. Propositional models, however, offered a new twist to the network idea. Instead of concept nodes comprising the basic unit of knowledge that is stored in memory, propositional models take this basic unit to be the proposition (Anderson & Bower, 1973). A proposition is a combination of concepts that has a subject and predicate. So, for example, instead of the concept *bird* representing a node in memory, the propositions "A bird has wings," "A bird flies," and "A bird has feathers" are stored.

There appears to be some psychological reality to the notion of propositions, because subjects will take longer to read sentences containing many propositions than those containing few, even when the number of actual words is the same (Kintsch, 1974). In addition, recall tends to reflect propositional structure rather than sentence structure. For example, suppose you read the following sentence as part of a passage on shorebirds: "The blue heron, a tall bird with a long neck and long legs, can usually be found in the marshy areas near water." Asked to recall later what you had read, you would be unlikely to reproduce this sentence. Instead, you might recall some of the ideas, or propositions, expressed in it, such as: "The blue heron is a tall bird. It has long legs and a long neck. It lives near water." For this reason, propositions have been used as a measure of recall in some memory experiments (e.g., Royer & Cable, 1975; Royer & Perkins, 1977).

John R. Anderson has developed perhaps the most comprehensive network model of memory that emphasizes propositional structure. Known initially as ACT (adaptive control of thought) (Anderson, 1976), the model evolved to ACT* as Anderson (1983) distinguished between procedural and declarative knowledge and added a system for modeling the long-term store of procedural knowledge. He has revised the model again (Anderson, 1996; Schooler & Anderson, 1997) to make it more consistent with research on the neural structure of the brain and to more strongly emphasize the adaptive nature of cognition. Now known as ACT-R, Anderson's model is so global that Leahey and Harris (1997) fear it may be too complex to definitively test or falsify.

Parallel Distributed Processing (PDP) Models of LTM. Parallel processing is distinguished from serial processing in that multiple cognitive operations occur simultaneously as opposed to sequentially. In a sentence verification task such as "A blue heron is an animal," for example, serial processing dictates that the search would start at *blue heron* and proceed along the pathways connected to the concept, one pathway at a time. In parallel processing, however, the search task is distributed, so that all possible pathways would be searched at the same time.

As they evolved, network models such as Anderson's came to include the assumption of parallel processing, but this assumption is at the very core of PDP, or connectionist, models of long-term memory. With connectionist models, researchers seek to describe cognition at a behavioral level in terms of what is known about actual neural patterns in the brain.

The PDP Research Group pioneered the development of these models (McClelland, Rumelhart, and the PDP Research Group, 1986; McClelland, 1988, 1994; Rumelhart, 1995), which propose that the building blocks of memory are connections. These connections are subsymbolic in nature, which means that they do not correspond to meaningful bits of information like concept nodes or propositions do. Instead, the units are simple processing devices, and connections describe how the units interact with each other. They form a vast network across which processing is assumed to be distributed. When learning occurs, environmental input (or input from within the network) activates the connections among units, strengthening some connections while weakening others. It is these patterns of activation that represent concepts and principles or knowledge as we think of it. This means that knowledge is stored in the connections among processing units.

Bereiter (1991) offered a "rough physical analogy" for understanding how a connectionist network might operate:

> Imagine that in the middle of a bare room you have a pile of a hundred or more frisbees, which are connected among themselves by means of elastic bands that vary in thickness and length. On each wall is a clamp to which you fasten a

frisbee. Take any four frisbees and clamp one to each wall. There will be an oscillation set up as those four frisbees pull on the other ones, and those pull on each other. In time, the oscillations will cease, and the frisbee population will settle into a pattern that reflects an equilibrium among the tensions exerted by the elastic bands. (p. 12)

If one were to change which frisbees are clamped to the wall or the lengths or thicknesses of the bands connecting the frisbees, oscillation would reoccur and a new pattern would settle out.

Because connections among units are assumed to carry different weights of association, learning occurs in the continual adjustment of the weights. Moreover, since processing occurs in parallel, many different adjustments can take place simultaneously, and there can be continuous error adjustments as a function of new information.

Consider how a PDP model might account for the experiences of Wes and Marcy in The Mechanic and the Web Surfer. In Marcy's case, the units and connections representing her knowledge of car mechanics are likely to be neither extensive nor strong, but some are already stronger than others. It is probably safe to assume, for example, that connections related to steering are stronger than those related to tie rods. Marcy's conversation with Wes serves to activate and strengthen further some of those connections and perhaps introduces new connections (e.g., steering damper may be a new concept to her, although both "steering" and "damper" are familiar). When it comes to recalling the conversation later, then, the stronger connection weights associated with "steering" enable Marcy to remember that as the gist of what was said. Likewise, the very weak connection weights associated with "steering damper" are not enough to prompt its specific recall. A similar analysis could be applied to Wes and what he remembers about the Internet.

PDP models offer a number of advantages over the other models in terms of what they explain about human information processing. First, they seem to account well for the incremental nature of human learning. With constant readjustment of connection weights, they provide a more dynamic picture of human learning than has been suggested heretofore (Estes, 1988). PDP models also offer "for the first time a convenient way of incorporating goals into the dynamics of information processing systems" (Estes, 1988, p. 207). That is, connection weights in most PDP systems are adjusted to reduce disparity between their output and some target output, which may be viewed as a goal.

Finally, there is potential in PDP models to explain cognitive development (McClelland, 1988, 1995). Some knowledge, in terms of prewired connection weights, can be built into the network. Exploring different configurations of initial memory architecture may lead to breakthroughs in determining just how much of human memory is "hard-wired."

Estes (1988) sounded some cautionary notes, however, concerning the conclusions over the long term to which PDP models may lead. He cited the

lack of forthcoming evidence to support PDP models as a mirror of neural processes in the brain. He reminded us that there is little reason to believe a single processor model will be sufficient to model brain functions. After all, "the evolution of the brain has not yielded a machine of uniform design like a digital computer but rather a melange of systems and subsystems of different evolutionary ages" (Estes, 1988, p. 206). He concluded that the final test of any theory will come in the record of extended research that follows from it.

Table 3.2 presents a summary of the models that have been proposed to account for how learned information is represented in memory. To this point, however, only verbal and procedural information have been addressed. What about memory for information of a visual or olfactory nature?

TABLE 3.2 ***Summary of Models Proposed to Account for the Storage of Information in Long-Term Memory***

Proposed Model	*Characteristics*	*Data the Model Explains*	*Difficulties Faced by the Model*
Network model	Memory represented as a web of nodes (concepts) connected by relations between concepts	Individual differences in memory Swift recognition of class and property relationships (e.g., bird has wings)	Cannot explain typicality of concepts (e.g., faster to recognize canary than penguin as birds)
Feature comparison model	Memory represented as sets of concept features	Typicality of concepts and "fuzzy sets"	Unwieldy and fails to account for semantic flexibility
Propositional model	Memory represented as network of propositions	Memory for gist Procedural as well as declarative knowledge	May be too complex to definitively test or falsify
Connectionist or PDP model	Memory represented as connections among subsymbolic units of processing	Incremental, dynamic nature of learning Possibility of hard-wiring of memory in the brain	A single model may be insufficient to represent brain functions

Dual-Code Models of LTM. Ask anyone what imagery is, and the response is likely to be, "pictures in my mind." Does this mean that imaginal information is represented in some way different from verbal information? How do we account for the variety of imaginal information, especially since there is more to imagery than just visual representations? We can imagine the tune of a favorite song, or the feel of a kitten's fur against our skin—examples of auditory and tactile imagery, respectively. In the same way, it is possible to generate examples of olfactory imagery ("Is that a hot apple pie I smell?") as well as kinesthetic imagery, which is often used in relaxation training.

Despite our subjective impressions of imagery, not all psychologists have been convinced of its existence as a separate form of information storage (e.g., Pylyshyn, 1973). Some investigations of visual imagery, for example, have shown that people remember a picture's meaning, rather than its visual attributes (e.g., Bower, Karlin, & Dueck, 1975; Light & Berger, 1976). This supports a unitary view of visual and verbal coding, which means that information about pictures is assumed to be represented in the same way as verbal information.

Other research, however, has challenged the unitary view. In a series of experiments conducted by Shepard and his associates (reviewed in Shepard, 1978), subjects appeared to mentally rotate images of three-dimensional figures in order to find their match among sets of distractors. That is, the amount of time it took to find a match was directly related to the number of turns required to rotate the test figure to the position of its match. This result held true even when subjects were given verbal instructions so that they had to rely on information in memory to generate the images.

The superiority of memory for concrete words over abstract words also poses problems for a unitary view of memory representation. People find it much easier to remember words like *sailboat, apple,* and *zebra* when they appear on a list than words such as *liberty* and *justice* (see, for example, Paivio, Yuille, & Rogers, 1969). If a dual-code or dual-systems approach is taken, however, these results are easy to explain. According to the dual-systems view (Paivio, 1971, 1986, 1991), there are two systems of memory representation, one for verbal information and the other for nonverbal information. Thus, for input such as concrete words, two codes are possible. The meaning of the words can be represented by the verbal system, but images of the words can also be represented by the imaginal system. With two memories available at recall, as opposed to one for abstract words, subjects should remember concrete words better.

Exactly how the imaginal system operates to store visual or other imaginal information is not known, although dual-code theorists agree that mental images are not exact copies of visual displays. Images tend to be imprecise representations, with many details omitted, incomplete, or, in some cases, inaccurately recorded. They also require effort to maintain and have parts that fade in and out (Kosslyn, 1980). Think of someone you know well,

for example, and try to visualize that person's face. Does he or she wear eyeglasses, and can you remember what they look like? Chances are you may remember verbally whether your friend wears glasses and then try to reconstruct visually what he/she looks like.

Researchers assume a strong connection between the verbal and imaginal systems, and for this reason, directions to form images and visual aids to instruction are both likely to enhance learning of some verbal material. Kosslyn (1980) suggested that images may be important to learning in enabling learners to represent what is not depicted in the instruction and then to transform these representations to facilitate comprehension and problem solving. Visual aids can function in the same way, particularly for learners with poor verbal skills (cf. Levin, 1983).

Retrieval of Learned Information

Once information has been stored in long-term memory, no matter in what form, it can be retrieved for use, retained over time, or forgotten. The process of **retrieval** from long-term memory is relatively simple to understand. Previously learned information is brought back to mind, either for the purpose of understanding some new input or for making a response. Using previous knowledge to understand and learn new material has already been discussed as encoding. But making a response based on previous knowledge raises the question, What kind of response? Consider the two questions below. Which question is likely to be more difficult to answer?

1. What does the word *esoteric* mean?
2. Which of the following words is the best synonym for *esoteric*?
   - **a.** essential
   - **b.** mystical
   - **c.** terrific
   - **d.** evident

Clearly, the first question is harder than the second because it provides fewer clues as to what the answer might be. This distinction between cued and noncued retrieval is the same as the difference between recall and recognition. To recall information, learners must both generate an answer and then determine whether it correctly answers the question. In recognition, however, potential answers are already generated, and the learner must only recognize which one is correct.

Recall. In free recall situations, learners must retrieve previously stored information with no cues or hints to help them remember. Subjects in many memory experiments, for example, are exposed to target information and then told to "write down everything you can remember about what you just read." Similarly, instructors ask such recall questions on tests as, "Write an

essay about America's involvement in World War II," or "Describe the con-
nectionist view of human memory." Because there are no cues present to po-
tentially bias retrieval, the output of free recall is assumed to represent
accurately what is in memory. However, researchers have found that the
amount subjects recall under these conditions tends to be low. Providing
them with cues raises the overall amount subjects are able to remember.

Cued recall tasks, then, are those in which a hint or cue is provided to help learners remember the desired information. This happens, for example, when teachers add qualifiers to their essay questions, such as "Be sure to discuss the role Pearl Harbor played in changing America's war policies." Leahey and Harris (1997) also cited the example of an actor learning lines as a cued recall task. Each line serves as a cue for remembering the next line.

Recognition. Recognition, in contrast to recall, involves a set of pregenerated stimuli presented to learners for a decision or judgment. In some cases, learners are asked to determine whether the stimulus information has been seen before, as in old-new recognition tasks. Tasks of this nature are common in memory experiments, but are becoming increasingly popular for assessing reading comprehension (e.g., Royer et al., 1984; Royer, 1990, 1995). For example, students read a target passage and then complete a sentence verification test. On the sentence verification test are test sentences of four types: (1) an original sentence from the passage; (2) a paraphrase of the original sentence in which the words are changed but the meaning is retained; (3) a meaning-change sentence in which one or two words in the original sentence are replaced to alter its meaning; and (4) a distractor sentence, which is consistent with the gist of the passage but unrelated to the original sentence. Students who comprehended the passage should be able to recognize the original and paraphrase sentences as old and classify the meaning-change and distractor sentences as new. Those who fail to comprehend the meaning of the passage, on the other hand, are likely to think that the meaning-change and distractor sentences are old on the basis of their similarity to sentences in the passage.

Two factors appear to influence old-new recognition. The most obvious is the strength of the memory trace, in that stronger memories will be more accurately recognized than weaker memories. But regardless of the strength of a memory trace, a decision must still be made about its match to the test stimulus. Imagine, for example, that you are choosing drapes to match the color of your living room carpet. You must make a decision concerning a particular set of drapes from your memory of the carpet's color. Now consider two possible scenarios: (1) the drapes are inexpensive, and besides, you can return them if the color is a poor match, (2) the drapes are expensive, must be paid for in advance, and cannot be returned. In which scenario are you more likely to make a yes decision?

The second factor influencing yes-no or old-new recognition is a decision criterion based on the context surrounding the recognition task. High-risk conditions lead to a more stringent criterion than do low-risk conditions, even though the memory trace in both situations is equivalent in strength and match to the test stimulus.

Besides yes-no recognition, there is also forced-choice recognition as exemplified in multiple-choice tests. As before, memory strength plays a role in the decision to choose a particular answer. The decision criterion, however, is determined not only by risk conditions surrounding the overall task, but by the distractors in each test item. That is, a severe penalty for wrong answers will decrease guessing overall, even though, in a four-distractor item, the chances of getting an item right by pure guessing is 25 percent. But suppose, in question 2, you could eliminate two of the distractors immediately. This increases to 50 percent the chances of getting the answer right, high enough odds, perhaps, to offset the penalty. An obvious implication of this for test construction is to write distractors that have equal probability of being chosen if the learner is forced to guess.

Encoding Specificity. Regardless of expected response type, the process of retrieval can be greatly influenced by the cues available to learners at test time. Two different principles have been investigated by researchers that suggest a relationship between conditions at encoding and conditions at recall.

The **encoding specificity** principle states, in essence, that *whatever cues are used by a learner to facilitate encoding will also serve as the best retrieval cues* for that information at test time (Thomson & Tulving, 1970; Tulving & Thomson, 1973). To illustrate, Anderson and Ortony (1975) gave subjects the sentences, "The container held the apples" and "The container held the cola." What images come to mind when you read those sentences? Most likely, you encoded an apple basket and a cola bottle. In fact, Anderson and Ortony found that *basket* served as an effective retrieval cue for the first sentence but not the second, while *bottle* served as a good cue for the second sentence but not the first.

Retrieval, then, is very much influenced by the context of encoding. This suggests for instruction that many different contexts or examples may be important to discuss during the presentation of new concepts. In this way, students will have many cues available to assist in encoding that may later be used for recall. If new information is presented in only one context, students may not find sufficient cues in test questions to support retrieval of information that is actually in memory.

Related to encoding specificity is the concept of state-dependent learning. Some years ago, a study was conducted in which subjects learned lists of paired words in one situation and recalled the lists in a different situation (Bilodeau & Schlosberg, 1951). The situations differed in the rooms in which the sessions (whether learning or testing) took place, whether the subjects were standing or sitting, and the method of list presentation. Results indicated that recall was best for those who were instructed and tested in the same situation. When the instructional situation differed from the testing situation, recall suffered. More recent studies on the effects of drugs have suggested that these recall differences can be explained in terms of the subjects' state of mind during learning and testing. Information learned in a particular state of mind (e.g., free from the influence of alcohol or other drugs) will be remembered best in the same state of mind (Goodwin et al., 1969).

Bower (1981) has demonstrated a similar phenomenon with moods. Words learned under a happy mood were better recalled under a happy mood than a sad mood, and words learned under a sad mood were best recalled in that state. Bower argued that emotions, just like information, are coded in memory. And indeed it seems likely that chemical changes in the brain induced by drugs, strong emotions, and learning may all be similarly explained.

Forgetting

At some point, all theories of memory must address the phenomenon of forgetting. We all forget things, but we may do so for many different possible reasons. The most common explanations for forgetting are failure to encode, failure to retrieve, and interference.

Failure to encode simply means that *the information sought during retrieval was never learned in the first place.* Learners often have the illusion of knowing. Poor readers, for example, typically do not monitor their reading very well and so believe they have read and understood something when they have not done so. Learners with ineffective study strategies face the same problem. They tend to equate effort with learning rather than monitor the actual effects of their learning strategies. A student in one of my classes, for example, could not understand why she had achieved such a low score on one of the examinations. "But I studied for hours!" she wailed. When I asked how she had studied, she looked back at me blankly—by rereading her notes and the book, of course. Repetition can only go so far. Elaboration may have helped to ensure that course material was solidly encoded in memory.

The concept of encoding failure emphasizes once again the importance of having and activating relevant prior knowledge in learning. In The Mechanic and the Web Surfer scenario, consider what relevant knowledge either Wes or Marcy could bring to bear in their discussions of car mechanics and Internet browsing. It is possible that each could retrieve enough to comprehend the other and respond appropriately during the conversation but not enough to encode details of the conversation for retrieval at a later time.

Failure to retrieve information that has been encoded in memory is a second cause of forgetting and refers to *the inability to access previously learned information.* This is something like losing the directory to your computer's hard drive. The files are still there, but without the appropriate cues (i.e., file names), they cannot be accessed and retrieved. Issues of encoding specificity and state-dependent learning have obvious relevance here. The more cues

that are used in encoding, the more likely one or another of them will be available to facilitate retrieval. In addition, assuming the validity of the dual-code theory, the more often encoding cues are generated in both the verbal and imaginal systems, the more likely retrieval will be facilitated.

A common strategy for enhancing retrieval is note-taking (Gagné & Driscoll, 1988). This is sometimes known as an external retrieval strategy (Kiewra, 1985; Kiewra & Frank, 1988; Kiewra et al., 1991) because its product—notes—serves as memory storage external to the learner. Students who elaborate on their notes also tend to perform better than those who simply reread them (Peper & Mayer, 1978), in essence optimizing the effects of encoding together with external retrieval.

Finally, long before the development of information-processing theory, **interference** was proposed as a cause of forgetting, which meant that *other events or information got in the way of effective retrieval.* McGeoch (1932) described forgetting of verbal materials in terms of two major laws. According to the first, forgetting was considered to be a function of the similarity between the circumstances of learning and testing, much as encoding specificity accounts for retrieval and forgetting now. The second set forth the conditions of interference, i.e., that numerous events and competing information can interfere with the retrieval of target information. Moreover, interference can occur from information learned either before or after the to-be-remembered information is learned. For example, retroactive interference has occurred when you read this chapter, read the next chapter, and then have difficulty recalling information from this chapter. Later learning interferes with the recall of earlier learned material, particularly as practice on the later material increases. This makes sense when we consider that information learned later is more recent and thus probably yields stronger memory traces than information learned earlier.

It is also possible, however, for previous learning to interfere with later learning. This is known as proactive interference, and the degree of interference is related to the amount of practice on the original task. Take, for example, the case of a long-time tennis player trying to learn racquetball. Since both are racket sports, it seems reasonable to believe that knowing one would facilitate learning the other. Instead, the well-learned skill of swinging a tennis racket interferes with the recently learned response of swinging a racquetball racket. Many players will find themselves swinging with the entire arm, as in tennis, rather than with just the wrist.

Proactive interference of a kind has also been demonstrated in the learning and memory of verbal materials by aging adults. Rice and Meyer (1985) investigated so-called memory deficits among older adults. Results of some studies had indicated that older adults remember less from a prose passage than do younger adults. In the series of experiments Rice and Meyer conducted, however, they found no evidence to support a memory deficit. Instead, they found that older adults, because they had so much more experience and prior knowledge, tended to get caught up in the details of the passage (which prompted reminiscing) and lose sight of the main ideas they were to recall. In other words, proactive interference had occurred. When main ideas were signaled, however, the effects of the interference were averted, and older adults remembered just as much as younger readers.

In a review of studies conducted with aging adults, Fry (1992) reached similar conclusions, and he suggested several concrete ways in which practitioners can help older adults learn and remember. For instance, visual displays of how the subject matter is structured and concepts related can provide useful encoding and retrieval cues. Similarly, because problems in the learning and remembering of adults seem to be a function of declining speed rather than declining mental powers, allowing adults to work at their own pace is a desirable instructional strategy. Finally, like children, adults can be taught more effective strategies for encoding and retrieval (Fry, 1992).

There is no denying that memory failure can also be caused by other conditions, such as amnesia or Alzheimer's disease. These causes, however, have relatively little relevance to instruction and are therefore beyond the scope of this chapter.

Implications of CIP for Instruction

Take a moment to reflect on the stages and processes of cognitive information processing that have been discussed in this chapter. What might they imply for instructional strategies? Some suggestions have already been made, and an integrated model of instruction based on CIP is presented in Chapter 10. Nonetheless, three general recommendations are worth exploring here. These are:

- Providing organized instruction
- Arranging extensive and variable practice
- Enhancing learners' encoding and memory
- Enhancing learners' self-control of information processing

Providing Organized Instruction

The organization of instruction has long been of interest to researchers because people will try to impose some meaningful structure or organization on any new information in order to make sense of it. So if learners are supposed to understand new information in particular ways, then the instruction must be organized to help them do this. As discussed earlier in the chapter, instructional tactics such as signaling what information is important and drawing learners' attention to specific features of that information can facilitate selective attention and appropriate pattern recognition. To enhance encoding and retrieval, as well as counteract the effects of interference, other tactics are appropriate, such as using imagery and representing information in multiple ways.

Graphic representations have been particularly effective in facilitating encoding and memory storage of information. Beissner, Jonassen, and Grabowski (1994; see also Jonassen, Beissner, & Yacci, 1993) reviewed the use of graphic techniques in acquiring structural knowledge, which represents relationships between concepts in a content domain. They concluded that graphic techniques (such as semantic maps, concept maps, networking) analyze, elaborate, and integrate subject matter content, as well as illustrate concept relations. The result is enhanced structural knowledge on the part of the learner. The concept maps that introduce each chapter of this book are a good example of graphic representations. If designed well, they should assist you in organizing and understanding the concepts discussed.

Arranging Extensive and Variable Practice

"Practice makes perfect" is a dictum well known to most learners, and in fact, there is some truth to the saying. As noted earlier in the chapter, automaticity of basic skills is a desirable educational goal, and extensive practice is one of the ways to help achieve it. Behavioral theorists referred to overlearning, or practicing a skill until it is so habitual as to require very little conscious attention. As will be seen in the discussion of learning motor skills (see Chapter 10), the amount of practice is not the only important variable. The kind of practice also matters. (As motor learning theorists are apt to say, "Perfect practice makes perfect!")

As noted from the evidence on encoding specificity, if the context changes substantially from encoding to retrieval, learners' performance may be impaired. Providing a great deal of varied practice helps learners to attach multiple cues to what they are learning, so they are more likely to recall it at test time in a range of appropriate contexts.

Enhancing Learners' Encoding and Memory

Many students come to college lacking study skills that will help them be successful as learners in the post–high school environment. Often, the goals they are asked to achieve are sufficiently more difficult than what they experienced in high school to put them at risk for failing. To help these students become better learners, community colleges and universities offer a variety of courses and experiences aimed at enhancing learners' encoding and memory. The strategies that are taught in these courses come directly from research on CIP that has been discussed in this chapter, and although they are aimed at college students, they are by no means limited to this population. Elementary and secondary school teachers, as well as instructional designers and trainers, can help the learners with whom they work to improve their encoding skills and memory.

Table 3.3 displays some suggested strategies for helping learners to enhance encoding and memory, along with the CIP process or principle with which they are most related. Think of how you might have used these strategies effectively in your own learning or how you might employ them with learners.

Enhancing Learners' Self-Control of Information Processing

When we shift the focus from instruction to learners, different aspects of information processing become prominent, suggesting different sorts of instructional implications. Earlier in the chapter, executive control processes were mentioned that enable the learner to modify information flow within and between components of the memory system. These processes have been investigated under the rubric of metacognition (Flavell, 1979; Brown, 1980;

TABLE 3.3 *Some Strategies for Enhancing Encoding and Memory*

Suggested Strategy	*Corresponding CIP Process*
Listen actively and pay attention to cues signaling what is important.	Selective attention
Encode information in more than one way and more than one mode. Use acronyms and imagery.	Dual code, multiple memory connections
Break down complex information into manageable parts.	Chunking
Elaborate on new information with examples that are meaningful to you.	Elaboration in encoding
Read actively. Make the information personal by relating it to your own life.	Elaboration in encoding
Take notes in your own words; don't just write it down verbatim.	Elaboration in encoding
Overlearn the material. Keep practicing even after you got them all right.	Rehearsal, automaticity
Review your class notes the same day that you take them.	Forgetting curve (Ebbinghaus)
Learn information in a similar way to what it needs to be recalled.	Encoding specificity
Avoid alcohol, caffeine, nicotine, or medications that might cause drowsiness during learning.	State dependent learning

Duell, 1986). **Metacognition** refers to *one's awareness of thinking and the self-regulatory behavior* (also known as conditional knowledge [cf. Prawatt, 1989]) *that accompanies this awareness.*

> In the course of learning and problem solving, representative kinds of regulatory performance include: knowing when or what one knows or does not know; predicting the correctness or outcome of one's performance; planning ahead and efficiently apportioning the outcomes of one's cognitive resources and one's time; and checking and monitoring the outcomes of one's solution or attempt to learn. (Gagné & Glaser, 1987, p. 75)

What is currently known about metacognitive skills and their acquisition goes well beyond the scope of this chapter, and the interested reader is referred to Derry and Murphy (1986) and Duell (1986) for their excellent reviews on the topic. Research results generally indicate, however, that metacognitive ability depends on person variables, task variables, strategy variables, and the interaction among all three (Duell, 1986).

With respect to person variables, older learners seem to have a better understanding of their memory abilities and limitations than do younger learners. Although students of all ages appear capable of learning various memory strategies, older learners are more planful and purposeful in their use of these strategies. Additionally, there is evidence that learning-disabled children are less efficient and less planful than normal children (Torgeson, 1977). This suggests that instructors should frequently remind younger and less planful learners when and how to use memory strategies.

Task variables refer to differences in instructional content that influence use of metacognitive strategies. For example, information that is new to learners will be approached with quite general learning strategies. As learners become more proficient in a subject or if the material they are to learn relates to a subject they know quite well already, they employ more domain-specific strategies (Gagné & Driscoll, 1988). For instructors to use or suggest the use of particular strategies, then, they should have some idea as to how much their students already know about the material to be learned.

Finally, strategy variables have to do with the metacognitive strategies themselves, the various ways in which learners may go about encoding, storing, and retrieving information. Some strategies are so simple that learners can acquire them easily by being told what to do. Breaking a complex or long learning task into manageable segments is one example. Other strategies, however, require extensive practice before they can be used easily and effectively. Taking notes or self-questioning with inferential questions may be examples of this type.

Educators generally agree on the importance of self-regulatory skills in learning, as will be especially evident in Chapter 9. Successful learners seem to acquire and refine these skills throughout their school and learning history. But what about the less successful and less proficient learners? Teaching learners to assume an active and purposeful role in their own learning has been a growing concern among instructors and researchers alike. Programs now exist to train students in metacognitive or study skills (e.g., Weinstein, 1982; Feuerstein et al., 1980; Dansereau et al., 1979; De Bono, 1985; Wang & Palincsar, 1989). Some are aimed at college students, others at younger learners. Some concentrate on domain-specific skills pertaining to a particular subject, such as reading comprehension; others train more general strategies that may be useful across a broad range of tasks. And some programs are embedded within school curricula, while others exist as separate, study skills courses.

Despite the variety among these programs, those that are effective seem to have at least two criteria in common. First, students must have a base of prior knowledge that may be related to the strategies they are learning. Domain-specific strategies, in particular, are virtually useless when students know little about the subject to which they pertain. Second, students must know when and why various self-regulatory strategies may be effectively employed (e.g., Pressley, Borkowski, & O'Sullivan, 1984; Prawatt, 1989; Sawyer, Graham, & Harris, 1992). Knowing how to be planful is not enough to guarantee that one will be planful. Having such conditional knowledge does not guarantee that one will always use it. But realizing when and why such behavior will be useful in furthering learning goals helps to motivate students to engage in metacognitive, self-regulatory ways.

Conclusion

As noted in the previous chapter, B. F. Skinner continued to argue against the necessity for inventing mental fictions to account for learning. At first, Roediger (1980) seemed to side with Skinner when he pointed out the proliferation of mental entities in current models of human memory and questioned what we have really learned from them. His conclusion, however, was not that mental constructs are useless, but that we should be cautious in what we take them to mean about learning and memory.

> Advances in theories of human memory parallel, and perhaps depend on, advances in technology.... The information processing approach has been an important source of models and ideas, but the fate of its predecessors should serve to keep us humble concerning its eventual success.... Unless today's technology has somehow reached its ultimate development, and we can be certain it has not, then we have not reached the ultimate metaphor for the human mind either. (Roediger, 1980, p. 244)

Cognitive information-processing theorists have not been the only ones interested in learning and memory from a cognitive perspective. In Chapter 4, the ideas of educational psychologist David P. Ausubel will be presented, along with the similar ideas of schema theory and mental models. In Chapter 5, situated cognition, with its emphasis on the integration of declarative and procedural knowledge, will be explored.

A CIP Look at "Kermit and the Keyboard"

Let us consider some cognitive information-processing concepts that might be relevant in understanding and explaining Kermit's learning in this story. An information-processing analysis of the act of performing a song at the keyboard might go something like this.

Kermit must first attend to the printed page of a musical score (the input). To process its contents requires recognition of the symbols (reading music is a process similar to reading text) and relating this to what he already knows. For instance, he notes the signature, which tells him how many beats per measure, and the key, which indicates how many sharps or flats. This information is retrieved to assist him in organizing a response, which is pressing down each key as it corresponds to that indicated in the score.

Frequent rehearsal helps Kermit's playing to become more automatic and less fraught with mistakes. Using different voices and backgrounds enables Kermit to vary the encoding cues so that he learns to play the same song in different contexts. One might explain his persistent error in "House of the Rising Sun" as a consequence of encoding specificity. He makes this mistake only when a particular background is used, the same background with which he made the mistake in the first place.

Reading the keyboard manual could be, for Kermit, very much like Rosemary's experience of reading the computer manual in the scenario "A Tale of Two Readers." Highly unfamiliar and complex and difficult content can cause comprehension problems, which Kermit encounters.

CIP offers a useful perspective on the continuing development of Kermit's keyboarding skills, but behaviorism provides a better explanation of why Kermit spends 20 minutes practicing some days and an hour other days. However, like behaviorism, CIP offers no particular insights into Kermit's motivation to study the keyboard to begin with.