Chapter

5

Information Processing Theory

Cass Paquin, a middle school mathematics teacher, seemed sad when she met with her team members Don Jacks and Fran Killian.

Don: What's the matter, Cass? Things got you down?

Cass: They just don't get it. I can't get them to understand what a variable is. "*X*" is a mystery to them.

Fran: Yes, "*x*" is too abstract for kids.

Don: It's abstract to adults too. "*X*" is a letter of the alphabet, a symbol. I've had the same problem. Some seem to pick it up, but many don't.

Fran: In my master's program they teach that you have to make learning meaningful. People learn better when they can relate the new learning to something they know. "*X*" has no meaning in math. We need to change it to something the kids know.

Cass: Such as what—cookies?

Fran: Well, yes. Take your problem $4x + 7 = 15$. How about saying: 4 times how many cookies plus 7 cookies equals 15 cookies? Or use apples. Or both. That way the kids can relate "*x*" to something tangible—real. Then "*x*" won't just be something they memorize how to work with. They'll associate "*x*" with things that can take on different values, such as cookies and apples.

Don: That's a problem with a lot of math—it's too abstract. When kids are little, we use real objects to make it meaningful. We cut pies into pieces to illustrate fractions. Then when they get older we stop doing that and use abstract symbols most of the time. Sure, they have to know how to use those symbols, but we should try to make the concepts meaningful.

Cass: Yes. I've fallen into that trap—teach the material like it's in the book. I need to try to relate the concepts better to what the kids know and what makes sense to them.

Information processing theories focus on how people attend to environmental events, encode information to be learned and relate it to knowledge in memory, store new knowledge in memory, and retrieve it as needed (Shuell, 1986). The tenets of these theories are as follows: "Humans are processors of information. The mind is an information-processing system. Cognition is a series of mental processes. Learning is the acquisition of mental representations." (Mayer, 1996, p. 154)

Information processing is not the name of a single theory; it is a generic name applied to theoretical perspectives dealing with the sequence and execution of cognitive events. Although certain theories are discussed in this chapter, there is no one dominant theory, and some researchers do not support any of the current theories (Matlin, 2009). Given this situation, one might conclude that information processing lacks a clear identity. In part this may be due to its influence by advances in various domains including communications, technology, and neuroscience.

Much early information processing research was conducted in laboratories and dealt with phenomena such as eye movements, recognition and recall times, attention to stimuli, and interference in perception and memory. Subsequent research has explored learning, memory, problem solving, visual and auditory perception, cognitive development, and artificial intelligence. Despite a healthy research literature, information processing principles have not always lent themselves readily to school learning, curricular structure, and instructional design. This situation does not imply that information processing has little educational relevance, only that many potential applications are yet to be developed. Researchers increasingly are applying principles to educational settings involving such subjects as reading, mathematics, and science, and applications remain research priorities. The participants in the opening scenario are discussing meaningfulness, a key aspect of information processing.

This chapter initially discusses the assumptions of information processing and gives an overview of a prototypical two-store memory model. The bulk of the chapter is devoted to explicating the component processes of attention, perception, short-term (working) memory, and long-term memory (storage, retrieval, forgetting). Relevant historical material on verbal learning and Gestalt psychology is mentioned, along with alternative views involving levels of processing and of memory activation. Language comprehension is discussed, and the chapter concludes by addressing mental imagery and instructional applications.

When you finish studying this chapter, you should be able to do the following:

- Describe the major components of information processing: attention, perception, short-term (working) memory, long-term memory.
- Distinguish different views of attention, and explain how attention affects learning.
- Compare and contrast Gestalt and information processing theories of perception.
- Discuss the major forms of verbal learning research.
- Differentiate short- and long-term memory on the basis of capacity, duration, and component processes.
- Define propositions, and explain their role in encoding and retrieval of long-term memory information.
- Explain the major factors that influence encoding, retrieval, and forgetting.
- Discuss the major components of language comprehension.
- Explain the dual-code theory and apply it to mental imagery.
- Identify information processing principles inherent in instructional applications involving advance organizers, the conditions of learning, and cognitive load.

INFORMATION PROCESSING SYSTEM

Assumptions

Information processing theorists challenged the idea inherent in behaviorism (Chapter 3) that learning involves forming associations between stimuli and responses. Information processing theorists do not reject associations, because they postulate that forming associations between bits of knowledge helps to facilitate their acquisition and storage in memory. Rather, these theorists are less concerned with external conditions and focus more on internal (mental) processes that intervene between stimuli and responses. Learners are active seekers and processors of information. Unlike behaviorists who said that people respond when stimuli impinge on them, information processing theorists contend that people select and attend to features of the environment, transform and rehearse information, relate new information to previously acquired knowledge, and organize knowledge to make it meaningful (Mayer, 1996).

Information processing theories differ in their views on which cognitive processes are important and how they operate, but they share some common assumptions. One is that information processing occurs in stages that intervene between receiving a stimulus and producing a response. A corollary is that the form of information, or how it is represented mentally, differs depending on the stage. The stages are qualitatively different from one another.

Another assumption is that information processing is analogous to computer processing, at least metaphorically. The human system functions similar to a computer: It receives information, stores it in memory, and retrieves it as necessary. Cognitive processing is remarkably efficient; there is little waste or overlap. Researchers differ in how far they extend this analogy. For some, the computer analogy is nothing more than a metaphor. Others employ computers to simulate activities of humans. The field of *artificial intelligence* is concerned with programming computers to engage in human activities such as thinking, using language, and solving problems (Chapter 7).

Researchers also assume that information processing is involved in all cognitive activities: perceiving, rehearsing, thinking, problem solving, remembering, forgetting, and imaging (Farnham-Diggory, 1992; Matlin, 2009; Mayer, 1996; Shuell, 1986; Terry, 2009). Information processing extends beyond human learning as traditionally delineated. This chapter is concerned primarily with those information functions most germane to learning.

Two-Store (Dual) Memory Model

Figure 5.1 shows an information processing model that incorporates processing stages. Although this model is generic, it closely corresponds to the classic model proposed by Atkinson and Shiffrin (1968, 1971).

Information processing begins when a stimulus input (e.g., visual, auditory) impinges on one or more senses (e.g., hearing, sight, touch). The appropriate *sensory register* receives the input and holds it briefly in sensory form. It is here that *perception (pattern recognition)* occurs, which is the process of assigning meaning to a stimulus input. This typically does not involve naming because naming takes time and information stays in the sensory register for only a fraction of a second. Rather, perception involves matching an input to known information.

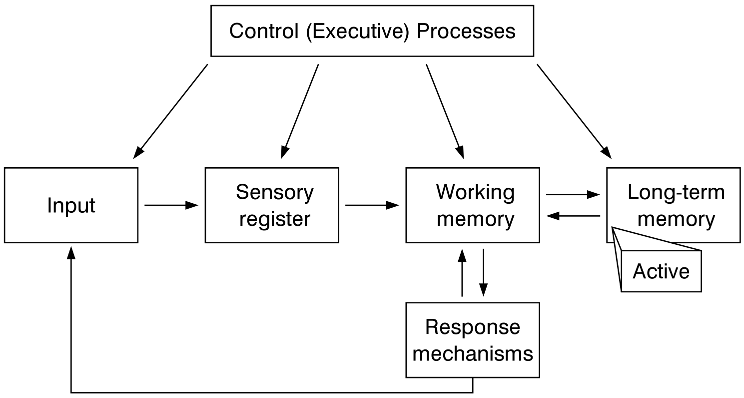


Figure 5.1
Information processing model of learning and memory.

The sensory register transfers information to *short-term memory (STM)*. STM is a *working memory (WM)* and corresponds roughly to awareness, or what one is conscious of at a given moment. WM is limited in capacity. Miller (1956) proposed that it holds seven plus or minus two units of information. A unit is a meaningful item: a letter, word, number, or common expression (e.g., "bread and butter"). WM also is limited in duration; for units to be retained in WM they must be rehearsed (repeated). Without rehearsal, information is lost after a few seconds.

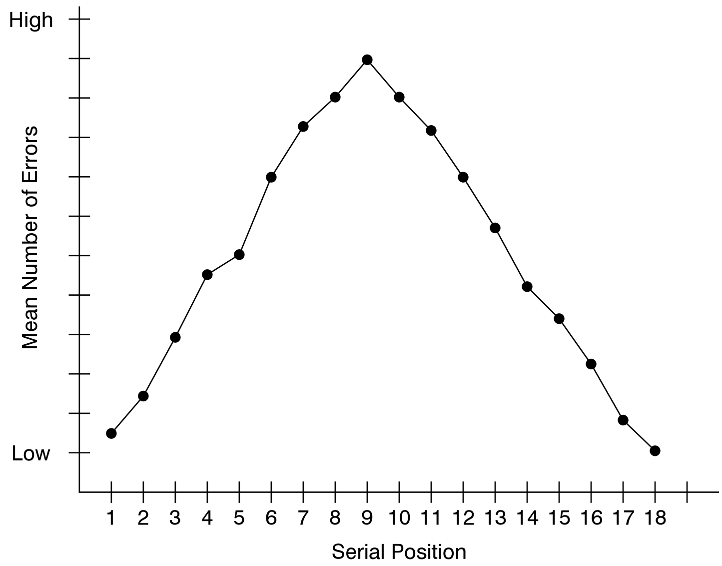
While information is in WM, related knowledge in *long-term memory (LTM)*, or permanent memory, is activated and placed in WM to be integrated with the new information. To name all the state capitals beginning with the letter *A*, students recall the names of states—perhaps by region of the country—and scan the names of their capital cities. When students who do not know the capital of Maryland learn "Annapolis," they can store it with "Maryland" in LTM.

It is debatable whether information is lost from LTM (i.e., forgotten). Some researchers contend that it can be, whereas others say that failure to recall reflects a lack of good retrieval cues rather than forgetting. If Sarah cannot recall her third-grade teacher's name (Mapleton), she might be able to if given the hint, "Think of trees." Regardless of theoretical perspective, researchers agree that information remains in LTM for a long time.

Control (executive) processes regulate the flow of information throughout the information processing system. Rehearsal is an important control process that occurs in WM. For verbal material, rehearsal takes the form of repeating information aloud or subvocally. Other control processes include coding (putting information into a meaningful context—an issue being discussed in the opening scenario), imaging (visually representing information), implementing decision rules, organizing information, monitoring level of understanding, and using retrieval, self-regulation, and motivational strategies. Control processes are discussed in this chapter and in Chapter 7.

The two-store model can account for many research results. One of the most consistent research findings is that when people have a list of items to learn, they tend to recall best the initial items (*primacy effect*) and the last items (*recency effect*), as portrayed in Figure 5.2. According to the two-store model, initial items receive the most rehearsal and

Figure 5.2
Serial position curve showing errors in recall as a function of item position.



are transferred to LTM, whereas the last items are still in WM at the time of recall. Middle items are recalled the poorest because they are no longer in WM at the time of recall (having been pushed out by subsequent items), they receive fewer rehearsals than initial items, and they are not properly stored in LTM.

Research suggests, however, that learning may be more complex than the basic two-store model stipulates (Baddeley, 1998). One problem is that this model does not fully specify how information moves from one store to the other. The control processes notion is plausible but vague. We might ask: Why do some inputs proceed from the sensory registers into WM and others do not? Which mechanisms decide that information has been rehearsed long enough and transfer it into LTM? How is information in LTM selected to be activated? Another concern is that this model seems best suited to handle verbal material. How nonverbal representation occurs with material that may not be readily verbalized, such as modern art and well-established skills, is not clear.

The model also is vague about what really is learned. Consider people learning word lists. With nonsense syllables, they have to learn the words themselves and the positions in which they appear. When they already know the words, they must only learn the positions; for example, "cat" appears in the fourth position, followed by "tree." People must take into account their purpose in learning and modify learning strategies accordingly. What mechanism controls these processes?

Whether all components of the system are used at all times is also an issue. WM is useful when people are acquiring knowledge and need to relate incoming information to knowledge in LTM. But we do many things automatically: get dressed, walk, ride a bicycle, respond to simple requests (e.g., "Do you have the time?"). For many adults, reading (decoding) and simple arithmetic computations are automatic processes that place little demand on cognitive processes. Such automatic processing may not require the operation of WM. How does automatic processing develop and what mechanisms govern it?

These and other issues not addressed well by the two-store model (e.g., the role of motivation in learning and the development of self-regulation) do not disprove the model; rather, they are issues to be addressed. Although the two-store model is the best-known example of information processing theory, many researchers do not fully accept it (Matlin, 2009; Nairne, 2002). Alternative theories covered in this chapter are levels (or depth) of processing and activation level, and the newer connectionism and parallel distributed processing (PDP) theories. Before components of the two-store model are described in greater detail, levels of processing and activation level theories are discussed (connectionism and PDP are covered later in this chapter).

Alternatives to the Two-Store Model

Levels (Depth) of Processing. *Levels (depth) of processing* theory conceptualizes memory according to the type of processing that information receives rather than its location (Craik, 1979; Craik & Lockhart, 1972; Craik & Tulving, 1975; Lockhart, Craik, & Jacoby, 1976). This view does not incorporate stages or structural components such as WM or LTM (Terry, 2009). Rather, different ways to process information (such as levels or depth at which it is processed) exist: physical (surface), acoustic (phonological, sound), semantic (meaning). These three levels are dimensional, with physical processing being the most superficial (such as "*x*" as a symbol devoid of meaning as discussed by the teachers in the introductory scenario) and semantic processing being the deepest. For example, suppose you are reading and the next word is *wren*. This word can be processed on a surface level (e.g., it is not capitalized), a phonological level (rhymes with *den*), or a semantic level (small bird). Each level represents a more elaborate (deeper) type of processing than the preceding level; processing the meaning of *wren* expands the information content of the item more than acoustic processing, which expands content more than surface-level processing.

These three levels seem conceptually similar to the sensory register, WM, and LTM of the two-store model. Both views contend that processing becomes more elaborate with succeeding stages or levels. The levels of processing model, however, does not assume that the three types of processing constitute stages. In levels of processing, one does not have to move to the next process to engage in more elaborate processing; depth of processing can vary within a level. *Wren* can receive low-level semantic processing (small bird) or more extensive semantic processing (its similarity to and difference from other birds).

Another difference between the two information processing models concerns the order of processing. The two-store model assumes information is processed first by the sensory register, then by WM, and finally by LTM. The levels of processing model does not make a sequential assumption. To be processed at the meaning level, information does not have to be first processed at the surface and sound levels (beyond what processing is required for information to be received) (Lockhart et al., 1976).

The two models also have different views of how type of processing affects memory. In levels of processing, the deeper the level at which an item is processed, the better the memory because the memory trace is more ingrained. The teachers in the opening scenario are concerned about how they can help students process algebraic information at a deeper level. Once an item is processed at a particular point within a level, additional processing at that point should not improve memory. In contrast, the two-store model contends that memory can be improved with additional processing of the same type. This model predicts that the more a list of items is rehearsed, the better it will be recalled.

Some research evidence supports levels of processing. Craik and Tulving (1975) presented individuals with words. As each word was presented, they were given a question to answer. The questions were designed to facilitate processing at a particular level. For surface processing, people were asked, "Is the word in capital letters?" For phonological processing they were asked, "Does the word rhyme with *train?*" For semantic processing, "Would the word fit in the sentence, 'He met a _____ in the street'?" The time people spent processing at the various levels was controlled. Their recall was best when items were processed at a semantic level, next best at a phonological level, and worst at a surface level. These results suggest that forgetting is more likely with shallow processing and is not due to loss of information from WM or LTM.

Levels of processing implies that student understanding is better when material is processed at deeper levels. Glover, Plake, Roberts, Zimmer, and Palmere (1981) found that asking students to paraphrase ideas while they read essays significantly enhanced recall compared with activities that did not draw on previous knowledge (e.g., identifying key words in the essays). Instructions to read slowly and carefully did not assist students during recall.

Despite these positive findings, levels of processing theory has problems. One concern is whether semantic processing always is deeper than the other levels. The sounds of some words (*kaput*) are at least as distinctive as their meanings ("ruined"). In fact, recall depends not only on level of processing but also on type of recall task. Morris, Bransford, and Franks (1977) found that, given a standard recall task, semantic coding produced better results than rhyming coding; however, given a recall task emphasizing rhyming, asking rhyming questions during coding produced better recall than semantic questions. Moscovitch and Craik (1976) proposed that deeper processing during learning results in a higher potential memory performance, but that potential will be realized only when conditions at retrieval match those during learning.

Another concern with levels of processing theory is whether additional processing at the same level produces better recall. Nelson (1977) gave participants one or two repetitions of each stimulus (word) processed at the same level. Two repetitions produced better recall, contrary to the levels of processing hypothesis. Other research shows that additional rehearsal of material facilitates retention and recall as well as automaticity of processing (Anderson, 1990; Jacoby, Bartz, & Evans, 1978).

A final issue concerns the nature of a level. Investigators have argued that the notion of depth is fuzzy, both in its definition and measurement (Terry, 2009). As a result, we do not know how processing at different levels affects learning and memory (Baddeley, 1978; Nelson, 1977). Time is a poor criterion of level because some surface processing (e.g., "Does the word have the following letter pattern: consonant-vowel-consonant-consonant-vowel-consonant?") can take longer than semantic processing ("Is it a type of bird?"). Neither is processing time within a given level indicative of deeper processing (Baddeley, 1978, 1998). A lack of clear understanding of levels (depth) limits the usefulness of this perspective.

Resolving these issues may require combining levels of processing with the two-store idea to produce a refined memory model. For example, information in WM might be related to knowledge in LTM superficially or more elaborately. Also, the two memory stores might include levels of processing within each store. Semantic coding in LTM may lead to a more extensive network of information and a more meaningful way to remember information than surface or phonological coding.

Activation Level. An alternative concept of memory, but one similar to the two-store and levels of processing models, contends that memory structures vary in their *activation level* (Anderson, 1990). In this view, we do not have separate memory structures but rather one memory with different activation states. Information may be in an active or inactive state. When active, the information can be accessed quickly. The active state is maintained as long as information is attended to. Without attention, the activation level will decay, in which case the information can be activated when the memory structure is reactivated (Collins & Loftus, 1975).

Active information can include information entering the information processing system and information that has been stored in memory (Baddeley, 1998). Regardless of the source, active information either is currently being processed or can be processed rapidly. Active material is roughly synonymous with WM, but the former category is broader than the latter. WM includes information in immediate consciousness, whereas active memory includes that information plus material that can be accessed easily. For example, if I am visiting Aunt Frieda and we are admiring her flower garden, that information is in WM, but other information associated with Aunt Frieda's yard (trees, shrubs, dog) may be in an active state.

Rehearsal allows information to be maintained in an active state (Anderson, 1990). As with working memory, only a limited number of memory structures can be active at a given time. As one's attention shifts, activation level changes.

We encounter the activation level idea again later in this chapter (i.e., Anderson's ACT theory) because the concept is critical for storage of information and its retrieval from memory. The basic notion involves *spreading activation*, which means that one memory structure may activate another structure adjacent (related) to it (Anderson, 1990). Activation spreads from active to inactive portions of memory. The level of activation depends on the strength of the path along which the activation spreads and on the number of competing (interfering) paths. Activation spread becomes more likely with increased practice, which strengthens structures, and less likely with length of retention interval as strength weakens.

One advantage of activation level theory is that it can explain retrieval of information from memory. By dispensing with the notion of separate memory stores, the model eliminates the potential problem of transferring information from one store to the other. STM (WM) is that part of memory that is currently active. Activation decays with the passage of time, unless rehearsal keeps the information activated (Nairne, 2002).

At the same time, the activation level model has not escaped the dual-store's problems because it too dichotomizes the information system (active-inactive). We also have the problem of the strength level needed for information to pass from one state to another. Thus, we intuitively know that information may be partially activated (e.g., a crossword item on the "tip of your tongue"—you know it but cannot recall it), so we might ask how much activation is needed for material to be considered active. These concerns notwithstanding, the activation level model offers important insights into the processing of information.

We now examine in greater depth the components of the two-store model: attention, perception, encoding, storage, and retrieval (Shuell, 1986). The next section discusses attention; perception, encoding, storage, and retrieval are addressed in subsequent sections.

ATTENTION

The word *attention* is heard often in educational settings. Teachers and parents complain that students do not pay attention to instructions or directions. (This does not seem to be the problem in the opening scenario; rather, the issue involves meaningfulness of processing.) Even high-achieving students do not always attend to instructionally relevant events. Sights, sounds, smells, tastes, and sensations bombard us; we cannot and should not attend to them all. Our attentional capabilities are limited; we can attend to a few things at once. Thus, attention can be construed as the process of selecting some of many potential inputs.

Alternatively, attention can refer to a limited human resource expended to accomplish one's goals and to mobilize and maintain cognitive processes (Grabe, 1986). Attention is not a bottleneck in the information processing system through which only so much information can pass. Rather, it describes a general limitation on the entire human information processing system.

Theories of Attention

Research has explored how people select inputs for attending. In *dichotic listening* tasks, people wear headphones and receive different messages in each ear. They are asked to "shadow" one message (report what they hear); most can do this quite well. Cherry (1953) wondered what happened to the unattended message. He found that listeners knew when it was present, whether it was a human voice or a noise, and when it changed from a male to a female voice. They typically did not know what the message was, what words were spoken, which language was being spoken, or whether words were repeated.

Broadbent (1958) proposed a model of attention known as *filter (bottleneck) theory.* In this view, incoming information from the environment is held briefly in a sensory system. Based on their physical characteristics, pieces of information are selected for further processing by the perceptual system. Information not acted on by the perceptual system is filtered out—not processed beyond the sensory system. Attention is selective because of the bottleneck—only some messages receive further processing. In dichotic listening studies, filter theory proposes that listeners select a channel based on their instructions. They know some details about the other message because the physical examination of information occurs prior to filtering.

Subsequent work by Treisman (1960, 1964) identified problems with filter theory. Treisman found that during dichotic listening experiments, listeners routinely shifted their attention between ears depending on the location of the message they were shadowing. If they were shadowing the message coming into their left ear, and if the message suddenly shifted to the right ear, they continued to shadow the original message and not the new message coming into the left ear. Selective attention depends not only on the physical location of the stimulus but also on its meaning.

Treisman (1992; Treisman & Gelade, 1980) proposed a *feature-integration* theory. Sometimes we distribute attention across many inputs, each of which receives low-level processing. At other times we focus on a particular input, which is more cognitively demanding. Rather than blocking out messages, attention simply makes them less salient than those being attended to. Information inputs initially are subjected to different tests for physical characteristics and content. Following this preliminary analysis, one input may be selected for attention.

Treisman's model is problematic in the sense that much analysis must precede attending to an input, which is puzzling because presumably the original analysis involves some attention. Norman (1976) proposed that all inputs are attended to in sufficient fashion to activate a portion of LTM. At that point, one input is selected for further attention based on the degree of activation, which depends on the context. An input is more likely to be attended to if it fits into the context established by prior inputs. While people read, for example, many outside stimuli impinge on their sensory system, yet they attend to the printed symbols.

In Norman's view, stimuli activate portions of LTM, but attention involves more complete activation. Neisser (1967) suggested that preattentive processes are involved in head and eye movements (e.g., refocusing attention) and in guided movements (e.g., walking, driving). Preattentive processes are automatic—people implement them without conscious mediation. In contrast, attentional processes are deliberate and require conscious activity. In support of this point, Logan (2002) postulated that attention and categorization occur together. As an object is attended to, it is categorized based on information in memory. Attention, categorization, and memory are three aspects of deliberate, conscious cognition. Researchers currently are exploring the neurophysiological processes (Chapter 2) involved in attention (Matlin, 2009).

Attention and Learning

Attention is a necessary prerequisite of learning. In learning to distinguish letters, a child learns the distinctive features: To distinguish *b* from *d*, students must attend to the position of the vertical line on the left or right side of the circle, not to the mere presence of a circle attached to a vertical line. To learn from the teacher, students must attend to the teacher's voice and ignore other sounds. To develop reading comprehension skills, students must attend to the printed words and ignore such irrelevancies as page size and color.

Attention is a limited resource; learners do not have unlimited amounts of it. Learners allocate attention to activities as a function of motivation and self-regulation (Kanfer & Ackerman, 1989; Kanfer & Kanfer, 1991). As skills become routine, information processing requires less conscious attention. In learning to work multiplication problems, students must attend to each step in the process and check their computations. Once students learn multiplication tables and the algorithm, working problems becomes automatic and is triggered by the input. Research shows that much cognitive skill processing becomes automatic (Phye, 1989).

Differences in the ability to control attention are associated with student age, hyperactivity, intelligence, and learning disabilities (Grabe, 1986). Attention deficits are associated with learning problems. Hyperactive students are characterized by excessive motor activity, distractibility, and low academic achievement. They have difficulty focusing and sustaining attention on academic material. They may be unable to block out irrelevant stimuli, which overloads their processing systems. Sustaining attention over time requires that students work in a strategic manner and monitor their level of understanding. Normal achievers and older children sustain attention better than do low achievers and younger learners on tasks requiring strategic processing (Short, Friebert, & Andrist, 1990).

Teachers can spot attentive students by noting their eye focus, their ability to begin working on cue (after directions are completed), and physical signs (e.g., handwriting) indicating they are engaged in work (Good & Brophy, 1984). But physical signs alone may not be sufficient; strict teachers can keep students sitting quietly even though students may not be engaged in class work.

Teachers can promote student attention to relevant material through the design of classroom activities (Application 5.1). Eye-catching displays or actions at the start of lessons engage student attention. Teachers who move around the classroom—especially

APPLICATION 5.1

Student Attention in the Classroom

Various practices help keep classrooms from becoming predictable and repetitive, which decreases attention. Teachers can vary their presentations, materials used, student activities, and personal qualities such as dress and mannerisms. Lesson formats for young children should be kept short. Teachers can sustain a high level of activity through student involvement and by moving about to check on student progress.

Kathy Stone might include the following activities in a language arts lesson in her third-grade class. As students begin each section of a teacher-directed exercise, they can point to the location on their papers or in their book. The way sections are introduced can be varied: Students can read together in small groups, individual students can read and be called on to explain, or she can introduce the section. The way students' answers are checked also can be varied: Students can use hand signals or respond in unison, or individual students can answer and explain their answers. As students independently complete the exercise, she moves about the room, checks students' progress, and assists those having difficulty learning or maintaining task focus.

A music teacher might increase student attention by using vocal exercises, singing certain selections, using instruments to complement the music, and adding movement to instruments. The teacher might combine activities or vary their sequence. Small tasks also can be varied to increase attention, such as the way a new music selection is introduced. The teacher might play the entire selection, then model by singing the selection, and then involve the students in the singing. Alternatively, for the last activity the teacher could divide the selection into parts, work on each of the small sections, and then combine these sections to complete the full selection.

Table 5.1
Suggestions for focusing and maintaining student attention.

Device	Implementation
Signals	Signal to students at the start of lessons or when they are to change activities.
Movement	Move while presenting material to the whole class. Move around the room while students are engaged in seat work.
Variety	Use different materials and teaching aids. Use gestures. Do not speak in a monotone.
Interest	Introduce lessons with stimulating material. Appeal to students' interests at other times during the lesson.
Questions	Ask students to explain a point in their own words. Stress that they are responsible for their own learning.

when students are engaged in seat work—help sustain student attention on the task. Other suggestions for focusing and maintaining student attention are given in Table 5.1.

Attention and Reading

A common research finding is that students are more likely to recall important text elements than less important ones (R. Anderson, 1982; Grabe, 1986). Good and poor readers locate important material and attend to it for longer periods (Ramsel & Grabe, 1983; Reynolds & Anderson, 1982). What distinguishes these readers is subsequent processing and comprehension. Perhaps poor readers, being more preoccupied with basic reading tasks (e.g., decoding), become distracted from important material and do not process it adequately for retention and retrieval. While attending to important material, good readers may be more apt to relate the information to what they know, make it meaningful, and rehearse it, all of which improve comprehension (Resnick, 1981).

The importance of text material can affect subsequent recall through differential attention (R. Anderson, 1982). Text elements apparently are processed at some minimal level so importance can be assessed. Based on this evaluation, the text element either is dismissed in favor of the next element (unimportant information) or receives additional attention (important information). Comprehension suffers when students do not pay adequate attention. Assuming attention is sufficient, the actual types of processing students engage in must differ to account for subsequent comprehension differences. Better readers may engage in much automatic processing initially and attend to information deemed important, whereas poorer readers might engage in automatic processing less often.

Hidi (1995) noted that attention is required during many phases of reading: processing orthographic features, extracting meanings, judging information for importance, and focusing on important information. This suggests that attentional demands vary considerably depending on the purpose of reading—for example, extracting details, comprehending, or new learning. Future research—especially neurophysiological—should help to clarify these issues (Chapter 2).

PERCEPTION

Perception (pattern recognition) refers to attaching meaning to environmental inputs received through the senses. For an input to be perceived, it must be held in one or more of the sensory registers and compared to knowledge in LTM. These registers and the comparison process are discussed in the next section.

Gestalt theory was an early cognitive view that challenged many assumptions of behaviorism. Although Gestalt theory no longer is viable, it offered important principles that are found in current conceptions of perception and learning. This theory is explained next, followed by a discussion of perception from an information processing perspective.

Gestalt Theory

The Gestalt movement began with a small group of psychologists in early twentieth-century Germany. In 1912, Max Wertheimer wrote an article on apparent motion. The article was significant among German psychologists but had no influence in the United States, where the Gestalt movement had not yet begun. The subsequent publication in English of Kurt Koffka's *The Growth of the Mind* (1924) and Wolfgang Köhler's *The Mentality of Apes* (1925) helped the Gestalt movement spread to the United States. Many Gestalt psychologists, including Wertheimer, Koffka, and Köhler, eventually emigrated to the United States, where they applied their ideas to psychological phenomena.

In a typical demonstration of the apparent motion perceptual phenomenon, two lines close together are exposed successively for a fraction of a second with a short time interval between each exposure. An observer sees not two lines but rather a single line moving from the line exposed first toward the line exposed second. The timing of the demonstration is critical. If the time interval between exposure of the two lines is too long, the observer sees the first line and then the second but no motion. If the interval is too short, the observer sees two lines side by side but no motion.

This apparent motion is known as the *phi phenomenon* and demonstrates that subjective experiences cannot be explained by referring to the objective elements involved. Observers perceive movement even though none occurs. Phenomenological experience (apparent motion) differs from sensory experience (exposure of lines). The attempt to explain this and related phenomena led Wertheimer to challenge psychological explanations of perception as the sum of one's sensory experiences because these explanations did not take into account the unique wholeness of perception.

Meaningfulness of Perception. Imagine a woman named Betty who is 5 feet tall. When we view Betty at a distance, our retinal image is much smaller than when we view Betty close up. Yet Betty is 5 feet tall and we know that regardless of how far away she is. Although the perception (retinal image) varies, the meaning of the image remains constant.

The German word *Gestalt* translates as "form," "figure," "shape," or "configuration." The essence of the *Gestalt psychology* is that objects or events are viewed as organized wholes (Köhler, 1947/1959). The basic organization involves a figure (what one focuses on) against a ground (the background). What is meaningful is the configuration, not the individual parts (Koffka, 1922). A tree is not a random collection of leaves, branches, roots, and trunk; it is a meaningful configuration of these elements. When viewing a tree, people typically do not focus on individual elements but rather on the whole. The human brain transforms objective reality into mental events organized as meaningful wholes. This capacity to view things as wholes is an inborn quality, although perception is modified by experience and training (Köhler, 1947/1959; Leeper, 1935).

Gestalt theory originally applied to perception, but when its European proponents came to the United States they found an emphasis on learning. Applying Gestalt ideas to learning was not difficult. In the Gestalt view, learning is a cognitive phenomenon involving reorganizing experiences into different perceptions of things, people, or events (Koffka, 1922, 1926). Much human learning is *insightful,* which means that the transformation from ignorance to knowledge occurs rapidly. When confronted with a problem, individuals figure out what is known and what needs to be determined. They then think about possible solutions. Insight occurs when people suddenly "see" how to solve the problem.

Gestalt theorists disagreed with Watson and other behaviorists about the role of consciousness (Chapter 3). In Gestalt theory, meaningful perception and insight occur only through conscious awareness. Gestalt psychologists also disputed the idea that complex phenomena can be broken into elementary parts. Behaviorists stressed associations—the whole is equal to the sum of the parts. Gestalt psychologists felt that the whole is meaningful and loses meaning when it is reduced to individual components. (In the opening scenario, "*x*" loses meaning unless it can be related to broader categories.) Instead, the whole is greater than the sum of its parts. Interestingly, Gestalt psychologists agreed with behaviorists in objecting to introspection, but for a different reason. Behaviorists viewed it as an attempt to study consciousness; Gestalt theorists felt it was inappropriate to modify perceptions to correspond to objective reality. People who used introspection tried to separate meaning from perception, whereas Gestalt psychologists believed that perception was meaningful.

Principles of Organization. Gestalt theory postulates that people use principles to organize their perceptions. Some of the most important principles are figure-ground relation, proximity, similarity, common direction, simplicity, and closure (Figure 5.3; Koffka, 1922; Köhler, 1926, 1947/1959).

The principle of *figure–ground relation* postulates that any perceptual field may be subdivided into a figure against a background. Such salient features as size, shape, color, and pitch distinguish a figure from its background. When figure and ground are ambiguous, perceivers may alternatively organize the sensory experience one way and then another (Figure 5.3a).

The principle of *proximity* states that elements in a perceptual field are viewed as belonging together according to their closeness to one another in space or time. Most people will view the lines in Figure 5.3b as three groups of three lines each, although other ways of perceiving this configuration are possible. This principle of proximity also is involved in the perception of speech. People hear (organize) speech as a series of words or phrases separated with pauses. When people hear unfamiliar speech sounds (e.g., foreign languages), they have difficulty discerning pauses.

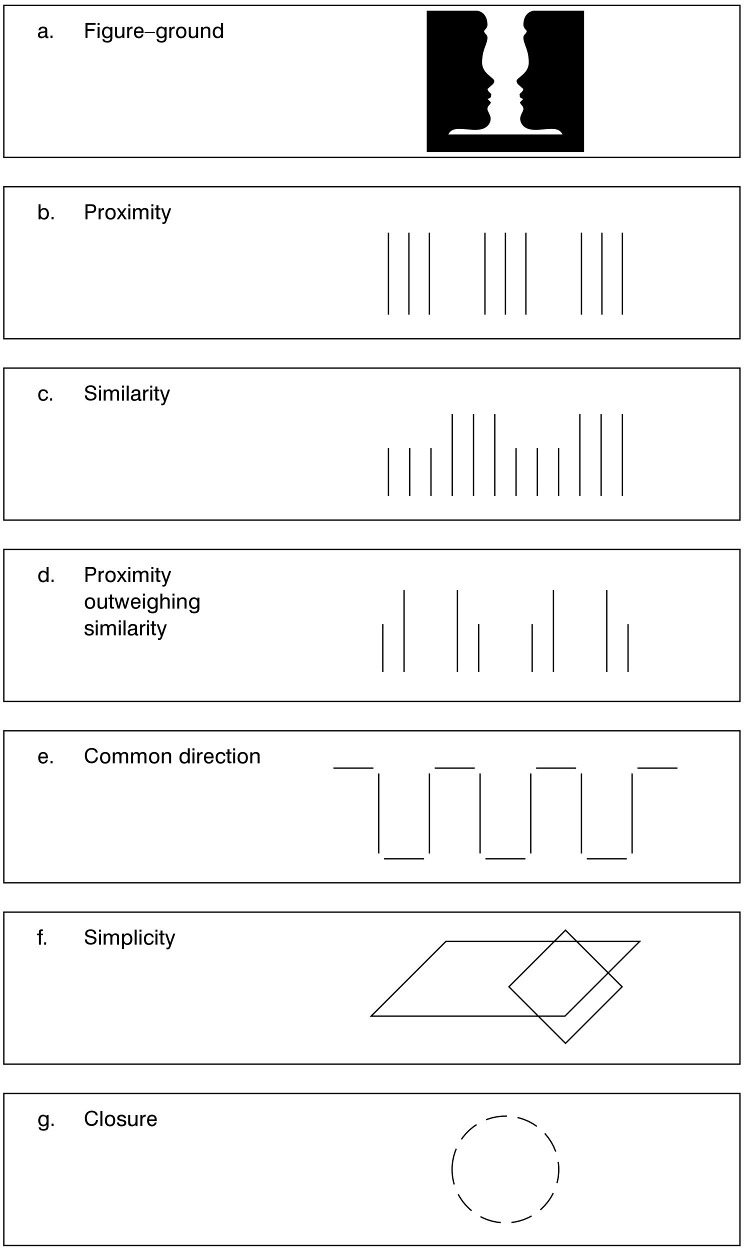


Figure 5.3
Examples of Gestalt principles.

The principle of *similarity* means that elements similar in aspects such as size or color are perceived as belonging together. Viewing Figure 5.3c, people tend to see a group of three short lines, followed by a group of three long lines, and so on. Proximity can outweigh similarity; when dissimilar stimuli are closer together than similar ones (Figure 5.3d), the perceptual field tends to be organized into four groups of two lines each.

The principle of *common direction* implies that elements appearing to constitute a pattern or flow in the same direction are perceived as a figure. The lines in Figure 5.3e are most likely to be perceived as forming a distinct pattern. The principle of common direction also applies to an alphabetic or numeric series in which one or more rules define the order of items. Thus, the next letter in the series *abdeghjk* is *m,* as determined by the rule: Beginning with the letter *a* and moving through the alphabet sequentially, list two letters and omit one.

The principle of *simplicity* states that people organize their perceptual fields in simple, regular features and tend to form good Gestalts comprising symmetry and regularity. This idea is captured by the German word *Pragnanz*, which roughly translated means "meaningfulness" or "precision." Individuals are most likely to see the visual patterns in Figure 5.3f as one geometrical pattern overlapping another rather than as several irregularly shaped geometric patterns. The principle of *closure* means that people fill in incomplete patterns or experiences. Despite the missing lines in the pattern shown in Figure 5.3g, people tend to complete the pattern and see a meaningful picture.

Many of the concepts embodied in Gestalt theory are relevant to our perceptions; however, Gestalt principles are quite general and do not address the actual mechanisms of perception. To say that individuals perceive similar items as belonging together does not explain how they perceive items as similar in the first place. Gestalt principles are illuminating but vague and not explanatory. Research does not support some of the Gestalt predictions. Kubovy and van den Berg (2008) found that the joint effect of proximity and similarity was equal to the sum of their separate effects, not greater than it as Gestalt theory predicts. Information processing principles, discussed next, are clearer and provide a better explanation of perception.

Sensory Registers

Environmental inputs are attended to and received through the senses: vision, hearing, touch, smell, and taste. Information processing theories contend that each sense has its own register that holds information briefly in the same form in which it is received (e.g., visual information is held in visual form, auditory information in auditory form). Information stays in the sensory register for only a fraction of a second. Some sensory input is transferred to WM for further processing. Other input is erased and replaced by new input. The sensory registers operate in parallel fashion because several senses can be engaged simultaneously and independently of one another. The two sensory memories that have been most extensively explored are *iconic* (vision) and *echoic* (hearing) (Neisser, 1967).

In a typical experiment to investigate iconic memory, a researcher presents learners with rows of letters briefly (e.g., 50 milliseconds) and asks them to report as many as they remember. They commonly report only four to five letters from an array. Early work by Sperling (1960) provided insight into iconic storage. Sperling presented learners with rows of letters, then cued them to report letters from a particular row. Sperling estimated that, after exposure to the array, they could recall about nine letters. Sensory memory could hold more information than was previously believed, but while participants were recalling letters, the traces of other letters quickly faded. Sperling also found that the more time between the end of a presentation of the array and the beginning of recall, the poorer was the recall. This finding supports the idea that forgetting involves *trace decay*, or the loss of a stimulus from a sensory register over time.

Researchers debate whether the icon is actually a memory store or a persisting image. Sakitt argued that the icon is located in the rods of the eye's retina (Sakitt, 1976; Sakitt & Long, 1979). The active role of the icon in perception is diminished (but not eliminated) if the icon is a physical structure, although not all researchers agree with Sakitt's position.

There is evidence for an echoic memory similar in function to iconic memory. Studies by Darwin, Turvey, and Crowder (1972) and by Moray, Bates, and Barnett (1965) yielded results comparable to Sperling's (1960). Research participants heard three or four sets of recordings simultaneously and then were asked to report one. Findings showed that echoic memory is capable of holding more information than can be recalled. Similar to iconic information, traces of echoic information rapidly decay following removal of stimuli. The echoic decay is not quite as rapid as the iconic, but periods beyond 2 seconds between cessation of stimulus presentation and onset of recall produce poorer recall.

LTM Comparisons

Perception occurs through bottom-up and top-down processing (Matlin, 2009). In *bottom-up processing,* physical properties of stimuli are received by sensory registers and that information is passed to WM for comparisons with information in LTM to assign meanings. Environmental inputs have tangible physical properties. Assuming normal color vision, everyone who looks at a yellow tennis ball will recognize it as a yellow object, but only those familiar with tennis will recognize it as a tennis ball. The types of information people have acquired account for the different meanings they assign to objects.

But perception is affected not only by objective characteristics but also by prior experiences and expectations. *Top-down processing* refers to the influence of our knowledge and beliefs on perception (Matlin, 2009). Motivational states also are important. Perception is affected by what we wish and hope to perceive. We often perceive what we expect and fail to perceive what we do not expect. Have you ever thought you heard your name spoken, only to realize that another name was being called? While waiting to meet a friend at a public place or to pick up an order in a restaurant, you may hear your name because you expect to hear it. Also, people may not perceive things whose appearance has changed or that occur out of context. You may not recognize co-workers you meet at the beach because you do not expect to see them dressed in beach attire. Top-down processing often occurs with ambiguous stimuli or those registered only briefly (e.g., a stimulus spotted in the "corner of the eye").

An information processing theory of perception is *template matching*, which holds that people store *templates*, or miniature copies of stimuli, in LTM. When they encounter a stimulus, they compare it with existing templates and identify the stimulus if a match is found. This view is appealing but problematic. People would have to carry around millions of templates in their heads to be able to recognize everyone and everything in their environment. Such a large stock would exceed the brain's capability. Template theory also does a poor job of accounting for stimulus variations. Chairs, for example, come in all sizes, shapes, colors, and designs; hundreds of templates would be needed just to perceive a chair.

The problems with templates can be solved by assuming that they can have some variation. Prototype theory addresses this. *Prototypes* are abstract forms that include the basic ingredients of stimuli (Matlin, 2009; Rosch, 1973). Prototypes are stored in LTM and are compared with encountered stimuli that are subsequently identified based on the prototype they match or resemble in form, smell, sound, and so on. Some research supports the existence of prototypes (Franks & Bransford, 1971; Posner & Keele, 1968; Rosch, 1973).

A major advantage of prototypes over templates is that each stimulus has only one prototype instead of countless variations; thus, identification of a stimulus should be easier because comparing it with several templates is not necessary. One concern with prototypes deals with the amount of acceptable variability of the stimuli, or how closely a stimulus must match a prototype to be identified as an instance of that prototype.

A variation of the prototype model involves *feature analysis* (Matlin, 2009). In this view, one learns the critical features of stimuli and stores these in LTM as images or verbal codes (Markman, 1999). When a stimulus enters the sensory register, its features are compared with memorial representations. If enough of the features match, the stimulus is identified. For a chair, the critical features may be legs, seat, and a back. Many other features (e.g., color, size) are irrelevant. Any exceptions to the basic features need to be learned (e.g., bleacher and beanbag chairs that have no legs). Unlike the prototype analysis, information stored in memory is not an abstract representation of a chair but rather includes its critical features. One advantage of feature analysis is that each stimulus does not have just one prototype, which partially addresses the concern about the amount of acceptable variability. There is empirical research support for feature analysis (Matlin, 2009).

Treisman (1992) proposed that perceiving an object establishes a temporary representation in an object file that collects, integrates, and revises information about its current characteristics. The contents of the file may be stored as an object token. For newly perceived objects, we try to match the token to a memorial representation (dictionary) of object types, which may or may not succeed. The next time the object appears, we retrieve the object token, which specifies its features and structure. The token will facilitate perception if all of the features match but may impair it if many do not match.

Regardless of how LTM comparisons are made, research supports the idea that perception depends on bottom-up and top-down processing (Anderson, 1980; Matlin, 2009; Resnick, 1985). In reading, for example, bottom-up processing analyzes features and builds a meaningful representation to identify stimuli. Beginning readers typically use bottom-up processing when they encounter letters and new words and attempt to sound them out. People also use bottom-up processing when experiencing unfamiliar stimuli (e.g., handwriting).

Reading would proceed slowly if all perception required analyzing features in detail. In top-down processing, individuals develop expectations regarding perception based on the context. Skilled readers build a mental representation of the context while reading and expect certain words and phrases in the text (Resnick, 1985). Effective top-down processing depends on extensive prior knowledge.

TWO-STORE MEMORY MODEL

The two-store (dual) memory model serves as our basic information processing perspective on learning and memory, although as noted earlier not all researchers accept this model (Matlin, 2009). Research on verbal learning is covered next to provide a historical backdrop.