CHAPTER ELEVEN

Imagery

Any single Circumstance of what we have formerly seen often raises up a whole Scene of Imagery.

—Joseph Addison

Scientific research on imagery is a relatively recent development, primarily because of the emphasis that behaviorists placed on overt behavior. Imagery, by its very nature, is internal, and asking subjects about their images seemed exactly like introspection. Because introspection as a method of inquiry had been shown to have a "dismal record of failure" in psychology (Bower & Clapper, 1989, p. 245), research on imagery was out of the provenance of respectable psychologists. In the late 1960s, this attitude began to change as several researchers demonstrated that imagery could be studied objectively without recourse to introspection.

Analog Versus Propositional Representations

One common definition of a *proposition* is that it is the smallest unit of knowledge that has a truth value. Propositions are usually represented as English sentences (at least in English-speaking countries), but they are not necessarily tied to language. Rather, a proposition is a basic unit of meaning that is either true or false and cannot be made any smaller without losing its truth value. For example, a proposition might be "Chicago is

Propositional

A is 5 miles north of B
C is 10 miles south of B
E is 20 miles west of B
E is 10 miles north of F
F is 20 miles west of C
C is 10 miles west of D

Analog

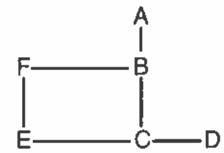

Figure 11.1 A comparison of a propositional and an analog way of representing the same information about the distances between six cities

north of Indianapolis" or "A robin has wings." No part of these propositions can be removed and still leave the proposition with a truth value.

An *analog* form of representation, by contrast, is one that preserves the structure of the original information in a more or less direct manner. A map is a good example of an analog form of representation; a map preserves the spatial structure of the actual physical environment but changes many other aspects of the information. It is often convenient to use a picture as an illustration of an analog form of representation. Just as propositions are not English sentences, analog representations are not pictures. What is important is that some aspect of the physical structure is represented in an accurate way in a more or less direct manner.

A good way to see the difference between an analog and a propositional form of representation is to compare the way the same information might look in each format. In Figure 11.1, both forms contain information about the relationships between six cities. In this example, most people would prefer the analog form to the propositional form because it is more readily useful for the purpose of navigating. For some forms of information, a propositional code might be more useful, whereas for other forms of information, an analog form might be more useful.

The central issue in the visual imagery literature is whether results from visual imagery experiments require two forms of representation or whether a propositional form can account for all the results. Almost all theorists agree that propositional forms are needed for representing knowledge; the question is whether analog forms are needed for imagery. At a general level, it is convenient to think of verbal codes (particularly knowledge and sentences) as being represented by propositions. In this chapter, we will see that it is equally convenient to think of visual codes (particularly visual images) as analog representations.

The Dual-Task Method

In 1968, Lee Brooks reported a series of experiments in which he had subjects perform two tasks simultaneously. The logic behind the *dual-task method* (also known as the *method of interference*) is that people can perform multiple tasks simultaneously as long as the resources required do not overlap. For example, many people can walk and chew gum at the same time; however, probably nobody can chew gum and whistle at the same time. The obvious reason is that the same resources are required in the latter task (moving the mouth, tongue, and lips), whereas different resources are required in the former case (legs versus mouth). Brooks used one task that was designed to require verbal processing and a

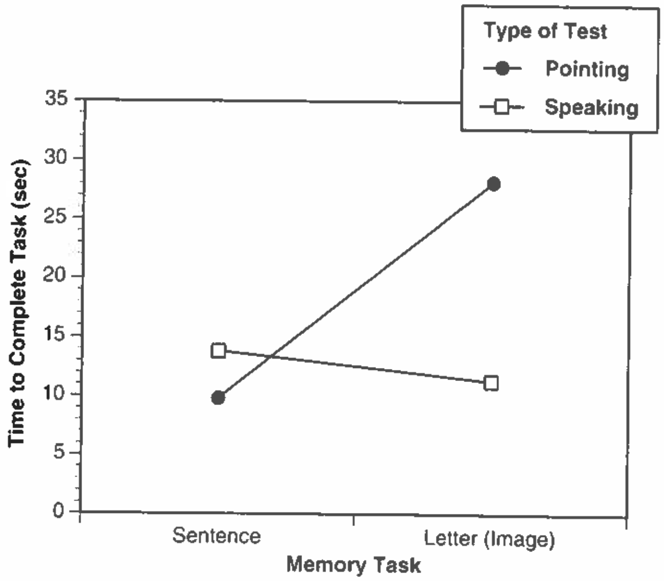

Figure 11.2 Time required to answer "yes" or "no" by speaking or pointing when the task involves processing either a sentence or an image. SOURCE: From "Spatial and Verbal Components of the Act of Recall," by L. R. Brooks, 1968, *Canadian Journal of Psychology, 22,* 349–368. Copyright © 1968 Canadian Psychological Association. Reprinted with permission.

second task that was designed to require spatial processing. He also devised two ways of responding: One required verbal answers, and the other required the subject to point to the correct answer.

Brooks (1968) used a verbal task and a verbal response, and a spatial task and a spatial response. In the verbal task, subjects were given sentences such as "A bird in the hand is not in the bush." The task was to identify each word as being a noun (respond "yes") or not a noun (respond "no"). Responses could be either verbal (saying the answer) or spatial (pointing to the answer). For the spatial response, subjects had to point to a *Y* or an *N* printed in irregularly staggered columns on a sheet of paper. In the imagery task, the subject was asked to imagine the outline of a letter, such as the letter *F,* and then to imagine a marker beginning at the lower left and moving around the outside. At each corner, the subjects were asked to indicate if it was an extreme top or bottom corner. For the letter *F,* correct responding would be "yes," "yes," "no," "no," and so on.

The results are shown in Figure 11.2. Note that the subject had to keep either the sentence or the image of the letter in memory while performing the response. When both the task and the response mode tapped the same mode of representation (letter form and pointing, sentence and speaking), responses were slower than when the task and response mode tapped different modes of representation (letter form and speaking, sentence and pointing). These results are consistent with the idea that there are two forms of representation—one corresponding roughly to a propositional or verbal format, and the other to an analog or spatial format. In his original report, Brooks (1968) included several other experiments to eliminate other possible explanations of the data.

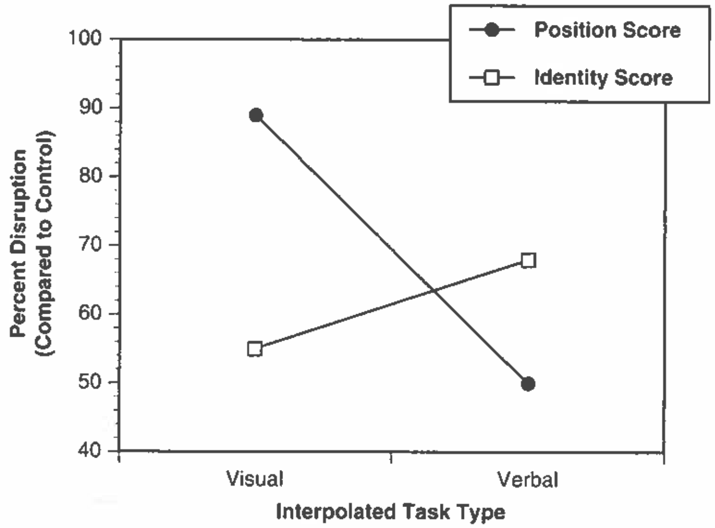

Figure 11.3 Percent disruption caused by an interpolated task compared to no interpolated task. Position scoring indicates that the correct cells were marked as having letters, regardless of whether the letter was correct. Identity scoring indicates that the correct letters were recalled regardless of which cell they were placed in. A disruption score of 0 would mean performance was the same as in the control group; larger scores mean more disruption. SOURCE: From "Selective Loss of Visual and Verbal Information in STM by Means of Visual and Verbal Interpolated Tasks," by K. den Heyer and B. Barrett, 1971, *Psychonomic Science, 25*, 100–102. Copyright © 1971. Reprinted by permission of Psychonomic Society Inc.

Den Heyer and Barrett (1971) replicated these results, using a paradigm in which both verbal and spatial aspects came from the same stimulus. Subjects briefly saw a grid in which several letters were placed; then, they were given one of three tasks. The first was to do nothing for 10 seconds (the control group), the second was to add several two-digit numbers, and the third was to compare three small grids and decide which was the odd one out. Then the memory test was given: The original grid was displayed again, and the subject's task was to enter the appropriate letters into the appropriate cells. Two different scores were calculated. The position score indicated how many cells were marked correctly as containing a letter, regardless of whether the letter entered was the correct one. The identity score indicated how many letters were recalled correctly, regardless of whether they were placed in the correct location.

The results are shown in Figure 11.3. The scores are shown in a measure called percent disruption. This score shows how much worse performance was in the two conditions with interpolated activity than in the control condition, in which there was no activity. A score of 0% would indicate equivalent performance; the larger the score, the more the interpolated activity disrupted performance. The pattern of results is the same as in Brooks's (1968) study: A visual interpolated task affected memory for location (measured by the position score) far more than it affected memory for specific letters (measured by the identity score). Conversely, a verbal interpolated task affected memory for specific letters far more than it affected memory for location.

Segal and Fusella (1970) found analogous results using auditory and visual stimuli: If subjects were asked to image a visual stimulus, they were worse at detecting a visual stimulus than an auditory one. Conversely, if they were asked to image an auditory stimulus, they were worse at detecting an auditory stimulus than a visual one. This is a modern demonstration of the Perky effect, named after Cheves Perky (1910). These findings are consistent with the idea of two forms of representation, propositional and analog. Much of the early theoretical work was done by Allan Paivio.

Paivio's Dual-Coding Theory

According to Paivio's (1971, 1986) *dual-coding theory*, words representing concrete objects can be encoded in two different ways, whereas words representing abstract objects can be encoded in only one way. For example, *dog* can be encoded using both a verbal code (the word itself) and a visual code (an image of a dog). The word *idea*, in contrast, probably has no visual code: What would an image of an idea look like? It may be possible to generate an image that is associated with *idea*, such as a lightbulb going on over a person's head, but that is not nearly as direct a coding as an image of a dog.

Because concrete words have two potential codes, compared to only one for abstract words, the dual-coding view predicts that concrete words should be remembered better; this is exactly what happens (Paivio, 1969). Furthermore, other things being equal, pictures are remembered better than words representing those pictures—the so-called picture *superiority effect* (Madigan, 1983; Paivio, 1971). The reason for the better performance is that two types of attributes are processed for concrete items; even if one of them fails to be useful, a second code is potentially retrievable. The benefit of concrete over abstract can

Experiment *Dual-Coding Theory*

Purpose: To demonstrate the advantage for concrete words over abstract words

Subjects: Twenty subjects are recommended.

Materials: Table 1 in the Appendix lists 80 abstract words and 80 concrete words of approximately equal word frequency and length. For each subject, create 10 lists of concrete words and 10 lists of abstract words; each list should have 8 different words. The first 10 lists should have 5 abstract and 5 concrete lists. You will also need an answer sheet on which subjects can write their responses.

Design: Because each subject experiences both conditions, concreteness is a within-subjects factor.

Procedure: Inform the subjects that in this experiment they will be asked to learn lists of words. Read each list at a rate of 1 word every 2 seconds. Then allow 30 seconds for free recall. Give a short break after 10 lists.

Instructions: "In this experiment, I will read you a list of words, and then I would like you to write down as many as you can remember on the answer sheet. Feel free to write the words in any order you like, and feel free to guess. After you have recalled each list, the next one will begin. Any questions?"

Scoring and Analysis: For each subject, count the number of concrete words recalled and the number of abstract words recalled. A paired *t* test can be used to verify that the means are different.

Optional Enhancements: Plot the data as a function of serial position.

Source: Based on an experiment by Paivio (1969).

be magnified if subjects are told to image the two objects interacting in some fashion. For example, Schnorr and Atkinson (1969) found that subjects could recall more than 80% of items when given imagery instructions, compared to about 40% when rote rehearsal instructions were used. However, imagery instructions do not have to be given for the advantage to be seen. Perhaps even more surprising, the advantage is seen in a free reconstruction-of-order test. In this test, a subject sees a list of words presented and is then given the list items in a new order. The task is to reconstruct the original presentation order. Even though all the words are given back to the subject, the concrete words are still remembered more accurately than the abstract words are (Neath, 1997).

There is a fair degree of similarity between Paivio's (1971, 1986) dual-code theory and the idea of the phonological loop and the visuo-spatial sketch pad in Baddeley's (1986; Baddeley & Hitch, 1974) working memory, as well as other proposals (Potter, 1979). The major difference between the two views is that whereas dual-coding assumes multiple memory codes that are modality specific, the other views include some more generic code common to both abstract and concrete information (Marschark, 1992).

Even with the success of dual-coding theory in accounting for these and other findings, much experimental work remained to be done to confirm the analog nature of visual imagery and to investigate further the properties of this system when engaged in more active processing of the image. Thus, studies turned away from examining the structure toward an emphasis on the function or dynamic properties of the analog form of representation.

Mental Rotation

If visual imagery is represented within an analog system that preserves the structure of the physical environment, then certain physical laws should be observed in imagery tasks. For example, if you hold a cube in your hand and rotate it 180°, the cube must first be rotated 1°, and then another 1°, and so forth. In other words, the cube has to pass through all intermediate orientations before completing the full 180°. This is a physical law in the real world. If the system representing visual information is preserving the spatial structure of the real world, as the studies by Brooks (1968) and den Heyer and Barrett (1971) suggest, then the same law should hold true when subjects image a cube rotating 180°.

CogLab Experiment
http://coglab.wadsworth.com/experiments/MentalRotation/

Note that this need not be the case. In fact, in a truly efficient system it is far quicker to simply flip the image 180° than to calculate each intervening orientation. When you ask a computer graphics program to rotate an object 180°, this is exactly what the computer does: a simple flip. To demonstrate that the imagery representation system is analog, one needs to show that an image of a rotated cube passes through all intermediate orientations.

This is exactly what Roger Shepard and his colleagues demonstrated (Cooper & Shepard, 1973; Shepard & Metzler, 1971; much of the work is reprinted in Shepard & Cooper, 1983). Shepard used objects that were made out of cubes (see Figure 11.4). Two such objects were shown, one on the left (the standard) and one on the right (the target). The subject was asked to decide if the target was the same as the standard, only rotated. Although the target was often not the same as the standard, the data of interest came

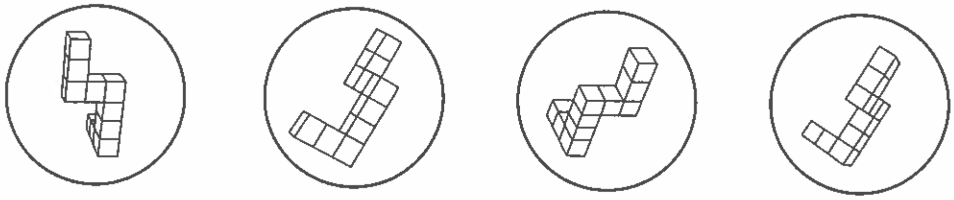

Figure 11.4 Two pairs of objects used in mental rotation studies. The subjects were shown the forms on the left and asked to decide if the form on the right was the same object, only rotated. SOURCE: From "Mental Rotation of Three-Dimensional Objects," by R. N. Shepard and J. Metzler, 1971, *Science, 171*, 701. Copyright © 1971 American Association for the Advancement of Science. Reprinted with permission.

when the target and standard were the same. In particular, Shepard was interested in how long it would take subjects to indicate that the objects were the same as a function of the degree of rotation. The results are shown in Figure 11.5.

The relationship between how far the object was rotated and the response time to decide that the two objects matched was linear, just as it would be in the real world: The further you rotate a cube in your hand, the longer it will take you, assuming the speed of rotation is constant. Notice also that it did not seem to matter which direction the rotation occurred, in either the picture plane (seeing an object rotate like the hands on a clock) or in the depth plane (seeing the top of the clock move toward you and then continue down as the bottom of the clock swings away from you and then up). Further work suggested that for a particular individual, the speed of rotation was constant (Finke, 1989).

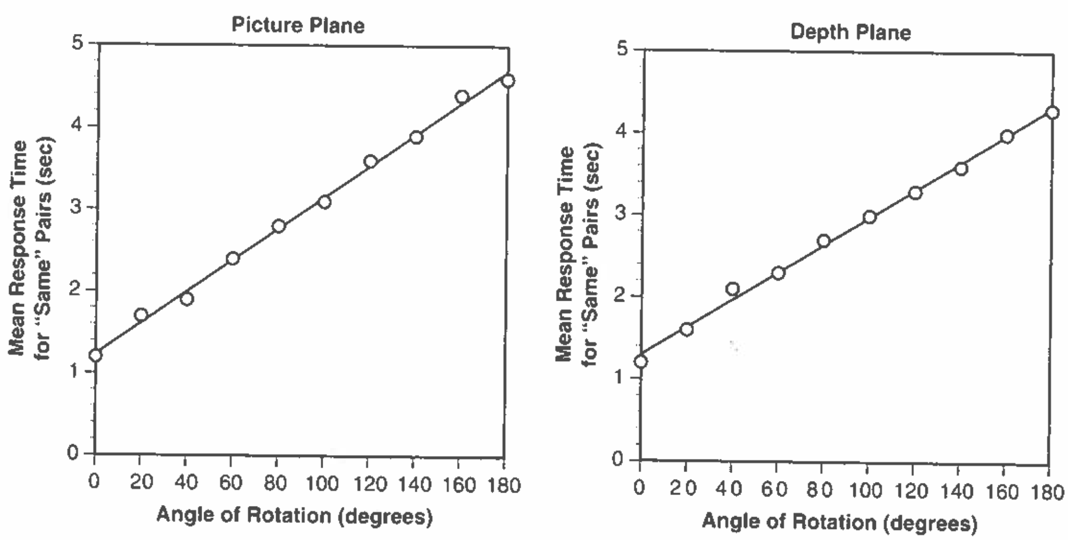

Figure 11.5 Mean response time to correctly say that the pair of objects matched as a function of the angle of rotation in degrees. The left panel shows objects that were rotated in the picture plane; the right panel shows objects that were rotated in depth. SOURCE: From "Mental Rotation of Three-Dimensional Objects," by R. N. Shepard and J. Metzler, 1971, *Science, 171*, 701–703. Copyright © 1971 American Association for the Advancement of Science. Reprinted with permission.

Of particular interest are the reaction times for 0° and 180°. The time taken to compare the two objects when both are in the same orientation can be taken as a baseline of how long the task and response take; any extra time can be attributed to the time needed to rotate. If subjects were simply flipping the object 180°, then the response time at this angle of rotation should be quite fast; the data clearly show that it was the slowest. This finding is consistent with the idea that the subjects were rotating through all intermediate positions.

Distinguishing Propositional from Analog Representation

Kosslyn (1976) reported an experiment consistent with the idea that both propositional and analog forms of representation are available. He selected features that were either highly associated or only marginally associated with a particular object. For example, claws are highly associated with the concept of a cat, whereas heads are not highly associated, even though cats have them. In addition, one of the features was small in relation to the main object, and the other large. Kosslyn's task was to have subjects verify statements such as "Cats have claws" and "Cats have heads." He predicted that if subjects were instructed to form an image, they should respond faster to the larger feature. In an image of a cat, the head is easier to "see" than the claws. However, if subjects were not instructed to use imagery, then the subjects should respond faster to the more highly associated item. As Figure 11.6 shows, this is exactly what happened.

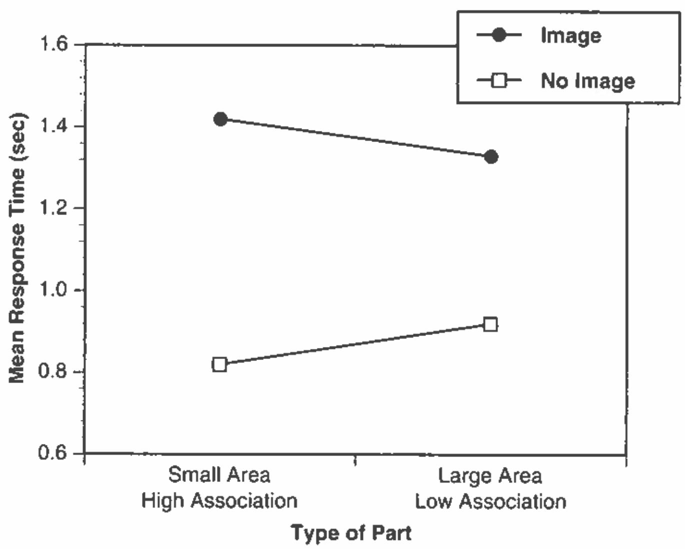

Figure 11.6 The time needed to verify a statement in which the part was either small but highly associated or large but had only a low association, as a function of instructions. Propositional theory predicts the results when no imagery instructions are given, but an analog theory better accounts for the results when imagery instructions are given.
SOURCE: 'From "Can Imagery Be Distinguished from Other Forms of Internal Representation?," by S. M. Kosslyn, 1976, *Memory & Cognition*, 4, p. 291–297. Copyright © 1976. Reprinted by permission of Psychonomic Society Inc.

Any theory of knowledge based on propositions would predict that subjects can retrieve information more quickly when the two concepts are highly associated than when they are less associated (Anderson, 1976). When no specific instructions are given, subjects will use a propositional form to get the answer because it is faster. Thus, subjects verified the statement "Cats have claws" faster than the statement "Cats have heads." However, when imagery instructions are given, subjects will generate an image and then answer the question based on the image. They will be able to retrieve information about a large part more quickly than a small part. Thus, subjects verified the statement "Cats have heads" faster than the statement "Cats have claws" because in the image, the head is larger than the claws.

Recall that Shepard and Metzler (1971) discovered a linear relationship between response time and degree of rotation. Kosslyn, Ball, and Reiser (1978) discovered that the same linear relationship holds for distance. They reported a study in which subjects first memorized a map of an island on which were seven objects, such as a hut, a swamp, a beach, and a well. These objects created 21 distances ranging from 2 cm to 19 cm. After subjects demonstrated that they had learned the map locations, the "mental scanning" phase of the experiment began. The subjects were asked to examine their image of the map and to focus on the named object, imagining a small speck hovering over the location. A second location was named, and the subject was asked to move the speck in a direct straight-line path and to press a button when the speck arrived at the destination. Kosslyn and his colleagues observed a linear relationship between distance scanned and response time; indeed, the correlation was 0.97.

Arguments Against Imagery

One of the most influential critics of the analog view of imagery is Zenon Pylyshyn (1973, 1979, 1981). Regardless of the outcome of the imagery debate, Pylyshyn's critiques have forced the research to become more rigorous, objective, and careful. Although his three objections are listed separately, they are not mutually exclusive, and they can operate simultaneously. Other arguments have also been offered against the need for a separate form of representation (Anderson, 1978).

One of Pylyshyn's objections is that *experimenter bias* may have determined experimental outcomes. It is well known that the expectations of the experimenter can influence how a subject responds in a particular situation (Rosenthal, 1966). For example, Intons-Peterson (1983) has shown that subtle cues provided by the experimenter can change the way the subject responds in an imagery experiment.

A second argument made by Pylyshyn is that subjects are using their tacit knowledge to guide their responses. *Tacit knowledge* is information people possess that reflects their beliefs about how things should work but that may not be verbalizable. (Tacit knowledge is often thought to be similar to implicit learning; see Chapter 7.) For example, because of interactions with the environment, people tacitly know that it takes more time to travel a long distance than a short distance at a given speed. They may not be able to formulate the equation rate × time = distance, but they are tacitly aware of this relationship. When asked to scan a visual image, subjects may assume that it should work just as physical motion does, and so their responses produce a linear relationship. The tacit knowledge point, then, says that the linear relationship between distance scanned and time to respond is due to the fact that subjects expect it to work this way.

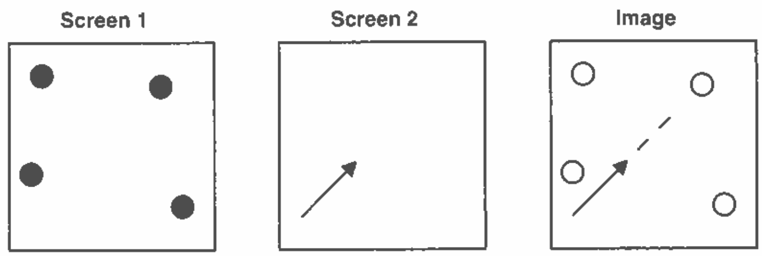

Figure 11.7 The subject saw the left panel for 5 seconds; then it disappeared. One second later, the middle panel was presented. The arrow could appear anywhere and could be in any orientation. Subjects reported deciding whether the arrow pointed to a dot by imagining they were scanning from the tip of the arrow along the direction specified until they hit a dot or a boundary. This strategy is represented in the right panel. SOURCE: Data from "Spontaneous Imagery Scanning in Mental Extrapolation," by R. A. Finke and S. Pinker, 1982, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8, 142–147. Copyright © 1982 American Psychological Association. Reprinted by permission of the author.

Pylyshyn's third argument suggests that task demands may be affecting the results. *Task demands* are similar to experimenter bias, except that it is the task itself, rather than any cues provided by the experimenter, that causes the subject to behave in one way rather than another. For example, when subjects are told that they are in an imagery experiment, the task is quite different from when no mention is made of the task.

To address Pylyshyn's first objection, Jolicoeur and Kosslyn (1985) conducted a replication of the Kosslyn et al. (1978) map study to see whether the experimenter's expectations affected the results. They told some of the experimenters to expect a U-shaped function between response time and distance scanned. Despite the experimenters' expectations, the usual linear function was found, and the Kosslyn et al. (1978) results were replicated. Although some results may be influenced by experimenter bias, the linear relationship between distance scanned and time to scan does not appear to be one of them.

An experiment by Finke and Pinker (1982) answered the last two criticisms. In this experiment, subjects saw a display that contained a random pattern of dots (see Figure 11.7). After 5 seconds, this display disappeared, followed 1 second later by a display that contained only an arrow. The task was to indicate whether the arrow pointed to any of the previously seen dots and to press a button to indicate yes or no.

Notice that in this experiment there is no mention of scanning, of forming an image, or even of the key variable—the distance between the tip of the arrow and the location of the dot. Nonetheless, the results were similar to those observed in other mental scanning studies: The farther the tip of the arrow was from the dot, the longer the response time. Kosslyn (1980) provides a model that addresses in detail how visual images might be constructed and manipulated.

Imagery and Perception

Despite the evidence supporting the analog view, there are still plausible theories of imagery that rely on propositional representations. For example, Humphreys and Bruce (1989) detail how both images and percepts may be built up from a common underlying propositional system. Regardless of which view ultimately explains the most data, both suggest great commonalities between the visual perception system and the visual imagery system.

Some of the most striking evidence comes from studies that tested patients with various visual deficits (reviewed by Farah, 1988). For example, Riddoch and Humphreys (1987) reported a case in which acquired cerebral color blindness (in which a patient loses the ability to perceive certain colors) was accompanied by a similar deficit in imagery for color. Levine, Warach, and Farah (1985) tested two patients, one with visual localization impairment and one with visual object identification impairment. The patient with object identification difficulties also had difficulties with tasks that required the formation of a visual image of an object. The patient with localization problems was unable to describe the relative locations of items in an image.

Another example concerns a person known as L. H. (Farah, Hammond, Levine, & Calvanio, 1988). L. H. had a perceptual deficit called agnosia; he had trouble identifying objects presented visually. He could see all the parts, such as lines, but could not assemble these parts together into whole objects. Consistent with the view that visual perception and visual imagery are related, L. H. had similar deficits with visual imagery.

A final example concerns patients with unilateral visual neglect. Patients with damage to the right parietal lobe often fail to detect stimuli presented in the left visual field, even though their basic perceptual processes are apparently normal (Posner, Walker, Friedrich, & Rafal, 1984). (Information from the left visual field goes to the right hemisphere, and information from the right visual field goes to the left hemisphere.) Patients with this disorder show a similar disorder in visual imagery tasks: Bisiach and his colleagues (Bisiach & Luzzatti, 1978; Bisiach, Luzzatti, & Perani, 1979) have demonstrated that when their patients were asked to imagine a famous square in Milan (with which the subjects were familiar), they failed to report landmarks that were in the left side of the scene. When the task was repeated, but the patients were asked to describe the square from a different side, the previously neglected information was now reported, because it was on the right side. Although unilateral neglect is usually attributable to an attentional disorder (e.g., Bisiach, 1999), the important point is that both vision and imagery are affected in a similar way. On the basis of these and other studies, Farah (1988) concluded that "for all of the types of selective visual deficits due to cortical lesions in which imagery has been examined, parallel imagery deficits have been observed" (p. 312). Farah (1989) offers one interpretation of how perception and imagery interact.

Another area of research relevant to the relationship between imagery and perception is called *memory psychophysics*. In traditional *psychophysics*, people perceive a stimulus and then make a judgment. In memory psychophysics, the same task is performed, but the judgment is made about a remembered stimulus. A standard technique is *magnitude estimation*, in which subjects rate a set of stimuli along some dimension such as size. The main restriction on the estimates is that they must preserve ratio properties. For example, if a subject sees one stimulus and rates it as a 10 and then sees another stimulus that is perceived to be twice as long, the subject should give it a 20, an estimate twice as large as the first. Typically, the estimated value (e) is related to the actual value (a) by a power function, $e = a^x$.

Kerst and Howard (1978) obtained magnitude estimations for distance (using centers of major U.S. cities) and area (of various states), and the exponents were calculated to be 1.04 and 0.79. Thus, subjects typically overestimated distance and underestimated area. Kerst and Howard also ran an imagery condition in which subjects first studied a map of the United States and then made judgments about distance and area from memory. The exponents in the imagery condition were calculated to be 1.10 and 0.60 for distance and area judgments, respectively. The important point is that the imagery exponents were the square of the perceptual exponents: $1.04 \times 1.04 = 1.08$, and $0.79 \times 0.79 = 0.62$. What this means, according to Kerst and Howard, is that in the imagery condition, the subjects were reperceiving the stimuli. There was one transformation during the encoding, when the actual distances were transformed with an exponent of 1.04. During the imagery phase of the experiment, the stimuli in memory, now an image, were reperceived and underwent a second transformation with an exponent of 1.04. These results have been replicated (Moyer, Bradley, Sorensen, Whiting, & Mansfield, 1978).

Algom and his colleagues (Algom, Wolf, & Bergman, 1985) conducted an extensive series of experiments on both perceptual and memory psychophysics that generally support the reperceptual hypothesis, at least for area judgments. The general methodology was to have subjects estimate the area of a rectangle given information about its height and width. This information could be provided entirely perceptually, entirely through memory, or in a mixture. Although the memory exponent was often close to the square of the perceptual exponent, there were several conditions under which this relationship did not hold. In all conditions, however, the memory exponent was smaller than the appropriate perceptual exponent. The researchers also found that whenever both perceptual and remembered information needed to be combined, the memory transformation dominated. Algom et al. concluded that their results support the analog view of imagery because of the "definitive evidence of second-order isomorphism between internal and external physical relations" (p. 469).

One further result consistent with the idea of some overlap between the visual imagery and visual perception systems comes from studies of visual imagery in congenitally blind subjects. These subjects, who have never seen anything, still show concreteness effects and effects attributable to visual imagery (De Beni & Cornoldi, 1988). Indeed, Kerr (1983) replicated the results of Kosslyn's map study (Kosslyn, Ball, & Reiser, 1978; Jolicoeur & Kosslyn, 1985) using blind subjects and a three-dimensional rendering of the map. These results are consistent with the idea that visual imagery is an analog form of representation that preserves spatial relationships. To the extent that the congenitally blind subjects had no deficit neurologically, then the standard visual imagery patterns should be observed.

Current Accounts of Imagery

There are two main theoretical accounts of imagery. One, usually referred to as visuo-spatial working memory (e.g., Logie, 1995), is related to Baddeley's (1986, 2000) working memory. The computational model of imagery, proposed by Kosslyn (1980, 1994), posits a system separate from memory. The two views have many similarities, but also some important differences. We focus primarily on the visuo-spatial account, as it is more integrated into theories of memory, whereas the computational model is more focused on the imagery processes of generation and transformation.

According to both views, images are generated or constructed sequentially (Kosslyn, 1994; Pearson, De Beni, & Cornoldi, 2001). This can be done either part by part or object by object, depending on what is being imaged. In general, a global image is generated first, followed by a second phase in which details are added that make the image more similar to visual experience. Images can be generated entirely from stored information, or can include information directly from perceptual experiences. Furthermore, there is more than one kind of visual image.

There is good evidence for a difference between visual imagery and spatial imagery (Farah et al., 1988). According to the visuo-spatial working memory view, the visuo-spatial sketch pad has a passive cache supported by an active spatial rehearsal mechanism called the visual scribe (Logie & Pearson, 1997). The architecture is similar to that seen in the phonological loop, where information decays within the phonological store unless rehearsed by the articulatory control process. In visuo-spatial working memory, information in the visual cache decays unless spatially rehearsed by the visual scribe. Information in the cache is also subject to retroactive interference from new items entering the cache.

The visual cache is used for retaining visual patterns, whereas the visual scribe is involved with retention of sequences of spatial movement. Evidence supporting this distinction comes from studies using the method of interference, described earlier in the chapter. For example, Logie and Marchetti (1991) found that memory for spatial patterns was disrupted by arm movements during the retention interval, whereas memory for purely visual information was not. Retention of visual information, but not spatial patterns, was disrupted when irrelevant pictures were shown during the retention interval.

Kosslyn (1994) suggests that the average duration of an image is 250 ms, which suggests that maintenance of the image is critical if the image is to be used in any kind of cognitive task. Furthermore, people seem to have difficulty maintaining an image for more than a few seconds (Cocude, Charlot, & Denis, 1997; Pazzaglia & Cornoldi, 1999). Kosslyn (1980, 1994) proposes that maintenance is an extension of generation, rather than a separate process. According to the visuo-spatial working memory view, some aspects of generation may contribute to maintenance, but other processes are also required (Logie, 1995; Pearson, De Beni, & Cornoldi, 2001)

Visuo-spatial working memory is not necessarily synonymous with visual imagery. Morton and Morris (1995) report a case study of a patient, M. G., who was impaired on typical mental imagery tasks, such as mental rotation and mental scanning, but was unimpaired on visuo-spatial working memory tasks involving the temporary maintenance of visual and spatial information. Additional evidence comes from the method of interference task. Although spatial tapping or arm movements can interfere with some mental imagery tasks, they do not always do so (Pearson, Logie, & Green, 1996; Salway & Logie, 1995). Further work is needed to understand fully the relationship between these processes and to develop a more complete theoretical account.

Real Versus Imagined Events

Given the similarities between visual imagery and visual perception, it might seem likely that the two events can sometimes be confused. The area of research that investigates how accurately people can distinguish real from imagined events is called *reality monitoring* and is related to the more general *source monitoring* (Mitchell & Johnson, 2000).

Johnson and Raye (1981) offer the most influential account of reality monitoring. According to their view, real (external) events can be distinguished from imagined (internal) events in two ways. The first is by an evaluation of qualitative attributes. Externally produced memories are likely to contain a rich representation of sensory attributes, including color, texture, intensity, and so forth. To the extent that a memory contains a rich encoding of sensory attributes and is highly detailed, it is likely to be an externally produced event. Internally generated events will lack some of this sensory information and will be less detailed and less integrated with contextual information. They often contain evidence of related cognitive processing, such as reasoning or decision making, that was necessary to generate the memory in the first place. For many memories, an evaluation of these details can determine the origin of the event, external or internal.

Many events, however—perhaps even a majority—will fall in a middle range. Although some features may indicate an external origin, a number of features may also indicate an internal origin. For these memories, Johnson and Raye (1981) propose a second kind of examination, based on plausibility and coherence. By *plausibility*, they mean that the memory must fit in with other knowledge about the world. A memory of flying by flapping one's arms is an extreme example of a lack of plausibility. *Coherence* refers to how well the memory fits in with other memories. For example, if supporting memories are available that make the memory of uncertain origin fit in with the grand scheme of things, then the uncertain memory is likely to be of a real event.

Note that these methods are not foolproof. Many sorts of memories are likely to fall into the middle category of uncertain. For example, Shepard (1984) points out that to be useful, intentions and plans need to be specific and detailed. Your memory for the plan is thus likely to have many of the attributes of an external memory, even though it was internally generated.

Many experiments support the general outline proposed by Johnson and Raye (1981). For example, Johnson, Raye, Wang, and Taylor (1979) divided subjects into two groups on the basis of their scores on a test that required them to remember details of pictures. During the experimental session, the subjects were shown pictures of common objects. Every so often, instead of a picture, the subjects saw the name of an object and were asked to create an image of the object. Johnson et al. varied the number of times each object was imaged. At the end of the session, the subjects were given an unexpected test in which they were asked to indicate the number of times they had seen a picture and to ignore the number of times they had imagined seeing it. Subjects who were deemed vivid imagers were less accurate than were subjects classified as less vivid imagers. According to Johnson et al., the rich imagery of the former group made it more difficult to distinguish instances of the imagined object from instances of the presented object.

The more similar the events in terms of their features, the more difficult it becomes to discriminate real from imagined. For example, Foley, Johnson, and Raye (1983) found that subjects were less accurate in determining whether they had spoken a word or imagined speaking the word than they were in deciding whether they had heard a word or imagined hearing the word. Similar results were reported by Anderson (1984).

Another quite systematic bias is known as the "it had to be you" effect (Anderson, 1984; Johnson & Raye, 1981). Here, the subjects appear to adopt the position that if they had really performed some action, it would be more memorable than if they had only imagined it. As a result, imagined events are more likely to be judged real than real events are to be judged imagined. Thus, memories of uncertain origins are more likely to be judged as originating externally than internally.

Further evidence comes from a study by Johnson, Kounios, and Reeder (1994), in which subjects saw some pictures and imagined others; at test, they received words and were asked to decide whether the object named had been seen, imagined, or neither.

Johnson et al. (1994) used a procedure whereby the subjects had to give a response as soon as they heard a tone. The tone could occur 300, 500, 900, or 1500 ms after the test item was displayed. At the short lags, Johnson et al. found that subjects were more accurate in determining "old" or "new" than they were in determining "seen" or "imagined." The researchers interpreted this finding as consistent with the idea that real and imagined events are made up of differing attributes and that the availability of these attributes can vary over time.

Johnson, Kahan, and Raye (1984) had subjects pair up and tell their partner about dreams they had actually dreamed, dreams they had made up, or dreams they had read about. Subjects were then given a discrimination test in which they had to decide whether a target event was one they had reported or one their partner had told them. Subjects had more difficulty with real dreams than with dreams they had read about or made up. In other words, they had difficulty recalling whether they had dreamed the dream or whether their partner had dreamed the dream. According to the reality-monitoring view, this is because dreams lack evidence of cognitive processes that are included in the memory when an event is imagined or read about. Based on this and other work, Johnson (1985) has suggested that memories generated under hypnosis may be more difficult to distinguish from real events (Dywan & Bowers, 1983) because of the extra vividness and degree of detail often present. Furthermore, such memories should lack evidence of cognitive operations because they were generated largely outside of conscious intent. According to this view, memories generated under hypnosis should be very difficult to distinguish from externally produced memories. In Chapter 12, we shall have more to say about the effects of hypnosis on memory.

Eidetic Imagery

Eidetic imagery is the technical name for *photographic memory*. The criteria for true photographic memory are very stringent. For example, a complex picture from *Alice's Adventures in Wonderland* might be shown for about 30 seconds; then, a variety of questions would be asked about specific details. The subject might be asked the number of stripes on the Cheshire Cat's tail or how many whiskers the cat had. To be truly photographic, the memory must preserve all the details that a photograph would.

The evaluation of eidetic imagery has remained constant between articles by Woodworth (1938, p. 45), Gray and Gummerman (1975), and Crowder (1992b). Eidetic images are more vivid and contain more detail than normal visual images do; they have a far longer duration than do afterimages or iconic memory; and they are found almost exclusively in preadolescent children.

Haber and Haber (1964) examined 150 preadolescent schoolchildren. The children were first shown afterimages (see Chapter 2) to get them used to talking about information not physically present. They were shown pictures for 30 seconds, and then the pictures were removed. The experimenters interviewed the children as they gazed at a blank sheet of cardboard where the picture had been. Although 84 demonstrated some evidence of imagery, only 12 (8% of the original sample) demonstrated eidetic imagery. These subjects were highly accurate and never used the past tense when answering. Furthermore, their eyes moved to the location on the blank where the target information had been displayed. Virtually no other children scanned the blank when answering.

Although the prevalence in preadolescent children is estimated at around 8% (Gray & Gummerman, 1975; Haber, 1979; Paivio & Cohen, 1979), the frequency in adults has been estimated as low as none in a million (Merrit, 1979). The notable exceptions to this are a report by Stromeyer and Psotka (1970) and one by Coltheart and Glick (1974); in the latter study, the subject could take a 10-word sentence and read off the letters backward. It is possible that the incidence of eidetic imagery is greater in elderly populations (Giray, Altkin, Vaught, & Roodin, 1985), but more data are needed on this point. Some authors assert that the famous Russian mnemonist S. used eidetic imagery. A careful reading of Luria's (1968) report, however, shows that although S. did make extensive use of visual imagery, he did not use eidetic imagery (see Chapter 15).

Many explanations have been offered for the drop in the prevalence of eidetic imagery from around 8% in preadolescent children to around 0% in normal adults, but none is universally accepted (Crowder, 1992b). One popular explanation is that as degree of literacy increases, reliance on the earlier pictorial form that gives rise to eidetic imagery decreases. However, Doob (1966) showed that degree of literacy, both within and between groups, is unrelated to prevalence of eidetic imagery. As Crowder (1992b) summarized the literature, "The most important conclusion about eidetic imagery is that it is a genuine phenomenon, capable of objective measurement and study" (pp. 155–156). Its cause and the reason for its disappearance and possible resurgence in older age are still unclear.

Other Forms of Imagery

So far, we have examined only visual imagery. More is known about both visual perception and visual imagery than about other forms of perception and imagery. Nonetheless, some research has focused on auditory imagery.

Although Perky (1910) included reports of auditory images and Seashore (1938) included a discussion in his text, much of the work on auditory imagery has been even more recent than that on visual imagery (see Reisberg, 1992). The nontemporal properties of a tone include pitch, loudness, and timbre. It is clear that the auditory image can contain information about pitch, although intensity is less reliably imaged (Intons-Peterson, 1992). Crowder (1989b) has used an interference demonstration to show that timbre can also be represented. Timbre is a potentially more interesting aspect, because it is more difficult for a subject to reproduce, physically, the unique sound of an oboe or a cello than it is to produce a sound of a particular frequency or intensity.

Crowder's (1989b) experiment involved two phases, the first demonstrating the perceptual effect and the second showing a similar effect in imagery. He presented two tones, and the subject's task was to decide whether they were the same or a different pitch. The pitches were either F, G, or A, and the timbres were of a flute, guitar, or trumpet. Subjects correctly said "same" more quickly when the pitch and timbre were identical than when only the pitches were identical and the timbres differed (see Figure 11.8). In the second phase, the tone was always a sine wave—a tone that essentially has no timbre—at one of the three pitches. After this tone was played, the subject was asked to imagine a tone of that pitch as produced by either a flute, a guitar, or a trumpet. The subject indicated when the image had been formed by pressing a button, and then the second note was played, this time by a flute, guitar, or trumpet. The task was to judge whether the pitch was the same or different, regardless of the timbre. Subjects again more quickly said "same" when the imagined and perceptual timbres matched than when they differed.

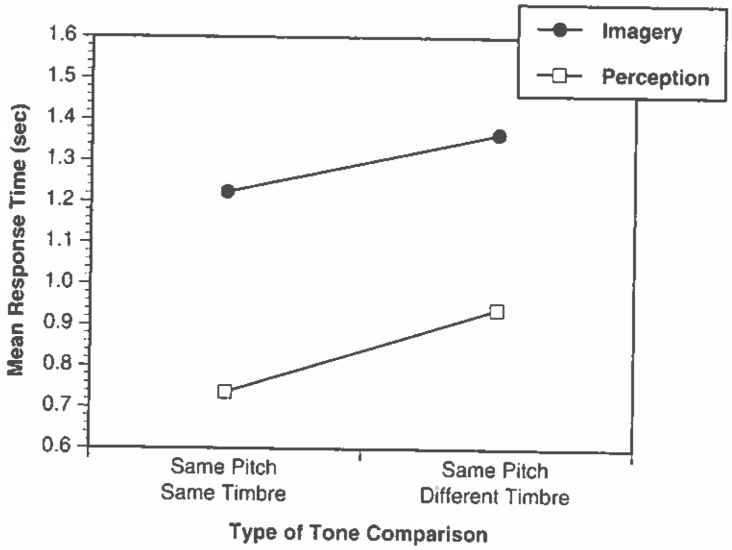

Figure 11.8 Mean response time to say that the second tone had the same pitch as the first tone when the timbre was the same or different. Subjects either heard (perception) or imaged (imagery) the timbre of the second tone.

SOURCE: From "Imagery for Musical Timbre," by R. G. Crowder, 1989, *Journal of Experimental Psychology: Human Perception and Performance*, 15, 472–478.

Voices can also be imaged (see, for example, Geiselman & Bjork, 1980; Nairne & Pusen, 1984), although this work has been less psychophysical than has research on auditory imagery for pitch, loudness, and timbre. For example, Geiselman and Glenny (1977) had subjects listen to a tape of either a male or female voice. They then saw words and were asked to imagine these words being pronounced by the voice they had just heard. At test, the subjects again heard words spoken in a male or female voice, but their task was simply to indicate whether the word spoken was a studied word. Subjects were more accurate for imagined words if the test word was presented in the same voice as imagined. MacKay (1992) reviewed work on a related area, termed *inner speech*.

There are relatively few theories of auditory imagery; indeed, Intons-Peterson (1992) states in her review, "I have not been able to find a single model of auditory imagery, per se" (p. 65). What models exist are either general models of imagery (Hebb, 1968) or simply adaptations of accounts of visual imagery.

Relatively little work has been done on tactile imagery. However, Marchant and Malloy (1984) did find that creating a tactile image facilitated memory for word pairs that were rated high in imagibility, just as using visual imagery facilitates memory for word pairs rated high in concreteness.

Crowder and Schab (1995) have reviewed imagery for odor, which appears somewhat different from visual and auditory imagery, at least based on the results of memory psychophysics experiments. Algom and Cain (1991) first trained subjects to associate a color with a particular odor. The odors were actually all the same (amyl acetate), but they differed in intensity. To ensure accurate memory, the subject sampled an odor and then gave its name (such as red). Twenty-four hours after learning, the magnitude estimation procedure began.

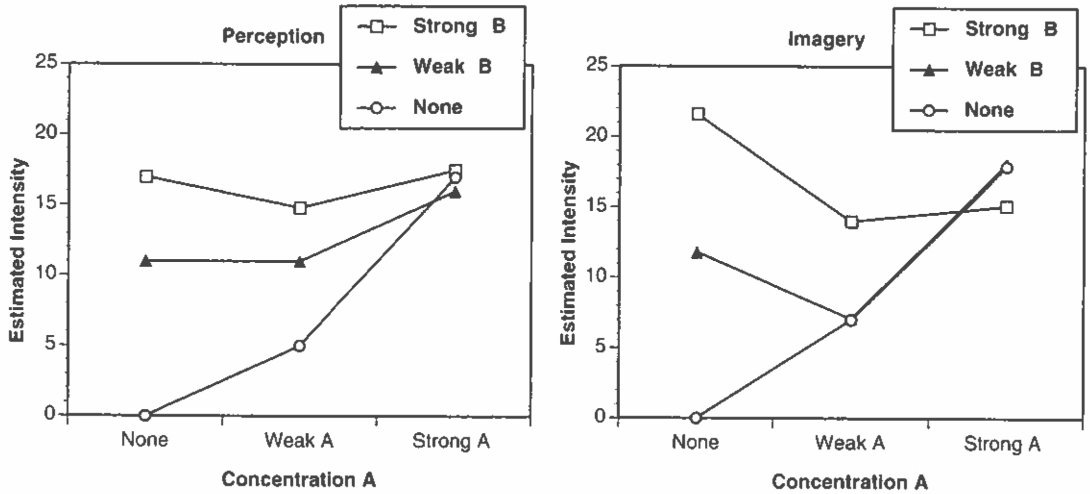

Figure 11.9 Estimated intensity when odors are perceived (left panel) and when odors are imaged (right panel). SOURCE: From "Remembered Odor and Mental Mixtures," by D. Algom and W. S. Cain, 1991, *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 1104–1119. Copyright © 1991 American Psychological Association. Reprinted with permission of the author.

Contrary to findings with visual stimuli, the exponents were almost identical for both perceptual and memory conditions. Thus, imagery for odors may be quite different from imagery for visual stimuli.

The best evidence for imagery for odors comes from a study in which subjects were able to combine and manipulate images of odors. Algom and Cain (1991) found that subjects performed very similarly when giving estimations of a mixture of two odors and when giving estimates of an imagined mixture. In this latter condition, the subjects first learned the individual odors as described above; in the psychophysical phase, they were asked to judge the intensity of mixing, for example, brown with red. Even though they had no physical experience with this mixture, their estimates were very similar to those given in the perceptual condition (see Figure 11.9).

Chapter Summary

One form of representation is a proposition, the smallest unit of knowledge that has a truth value. An analog form of representation is one that preserves the structure of the original information in a more or less direct manner. Solid evidence for the existence of two forms of representation first came from experiments that used the method of selective interference and from work on dual-coding theory, the idea that words can be represented in both a verbal and a visual code. Much of the research focused on eliminating other possible explanations—such as experimenter bias, tacit knowledge, and task demands—for results obtained in mental rotation and visual imagery experiments.

Although the relation between visual perception and visual imagery is still not completely understood, there are many intriguing parallels. Disruptions in visual imagery are often mirrored with similar disruptions to visual perception. Complicating this, however, is evidence that spatial imagery differs from visual imagery.

Reality monitoring refers to the ability to tell the difference between real events and imagined events. People are more likely to misremember an imagined event as real than a real event as imagined. Eidetic memory is the technical term for what most people refer to as photographic memory. About 5%–10% of preadolescent children seem to have this ability, but no adults do, with the possible exception of very elderly adults.

Just as people can create visual images, they can also create auditory images, tactile images, and odor images, although these have been the subjects of far less research than has visual imagery.