International Handbook of Psychology

edited by

KURT PAWLIK
AND
MARK R. ROSENZWEIG

SAGE Publications
London • Thousand Oaks • New Delhi
Contemplation of the marvels of human memory is one of the oldest topics in the Western tradition. Lucid descriptions of the puzzles provided by memory were included in the dialogues of Plato and the discourses of Aristotle. Philosophers provided 2,000 years of interesting theorizing and speculation before the methods of science were directed at the issue of how memory works. The starting point for scientific exploration of memory was the great experimental work begun by Hermann Ebbinghaus (1850–1909) in 1879, culminating in publication of Über das Gedächtnis in 1885 (with an English translation published in 1913 with the title, Memory: A contribution to experimental psychology). In this chapter we briefly summarize the modern understanding of human memory from the perspective of experimental psychologists. However, many different traditions exist among scholars and researchers studying memory, from humanistic traditions in fields such as history and literature to other scientific approaches from neurobiology, cognitive neuroscience, and animal learning and behavior. We make occasional reference to these other fields, but concentrate here on research from experimental psychology.

We begin the chapter with two historical landmarks of the field. Next we turn to some basic methods that are commonly used and some fundamental findings that have resulted from these methods. Following this section comes a brief consideration of theories of memory phenomena. We next include an experiment in some detail that provides a sample of research and how it is conducted. Finally, we consider some ways the psychology of memory has had an impact on applied issues outside the laboratory.

7.1 Two Approaches to the Study of Memory

Ebbinghaus’s scientific achievement was remarkable. He was the first person to empirically study memory, so he could not rely on the knowledge gained in past experiments to guide his own research. In light of such conditions,
one may have expected the first advances in the study of memory to have been small ones, perhaps a few poorly crafted, ill-controlled experiments leading to results that could not be replicated by later researchers. However, nothing could be further from the truth in this case. Ebbinghaus developed impeccable scientific methods and collected data that have proved reliable; all his main experimental results have been replicated.

Ebbinghaus developed his own materials, invented an ingenious method of measuring retention, and examined many variables that have potent effects on memory. The materials were three-letter nonsense syllables (e.g., ZAK, FER), which were intended to minimize the influence of past language skills on learning new information. He placed these syllables in long lists and tried to memorize them by reading aloud each item in time to a metronome so as to guarantee control over study time. After reading the list, he would try to recite it. If he did not succeed, he would read it and again attempt to recite it from memory. The measure of learning was then the number of these study/test trials (or the amount of time) that it would take to reach the criterion of one perfect recitation. This measure is called a trials-to-criterion measure of learning.

But how to measure retention at some later point, say a week later? Ebbinghaus could have tried to recall the list, but he memorized so many there was no way to keep them straight. Further, he realized that even if he could not recall the list, this failure might not indicate that all traces of it had vanished from memory. Ebbinghaus solved the problem by developing his ingenious relearning and savings method. When he wanted to test his retention at some later point, he would again present the list to himself and try to recite it, repeating the process until he could again correctly recall the entire list. The smaller number of trials needed to relearn the list than had been required to learn it on the first occasion provided a measure of savings. He expressed this difference as a function of the number of possible trials that could be saved multiplied by 100 to create a percentage savings score. So, for example, if he took 15 trials to learn the list on the first occasion, but only five trials to perform the same feat a week later, the savings score would be $15 - 5 = 15 \times 100$ for 66% savings, in this case. Of course, if the list took 15 trials to relearn, then the savings score would be zero (it took as many trials to relearn the list as it took to learn in the first place) and no measurable retention of the list would be obtained. The savings measure permits detection of retention when other measures such as recall or even recognition might fail. (However, even if savings is zero, that does not necessarily mean that no effect of the prior experience remains; it just means that this measure failed to detect it.)

Ebbinghaus (1885/1913) conducted several year's worth of research with his new methods and reported the results in his book. He showed some findings that are obvious—longer lists take more trials to learn than shorter lists—but many others which were less obvious. For example, presentations of lists that were spaced apart in time led to better long-term retention (greater savings) than did massed presentation of the lists, or when the same list was repeatedly read in succession. Besides his elegant results (all obtained with himself as the only subject!), Ebbinghaus also provided an excellent analysis of statistics, developed an early mathematical model (of forgetting), provided a good account of the problem of experimenter bias, and created what may have been the first example of competitive hypothesis testing in experimental psychology. Ebbinghaus started the scientific study of memory off on a sure footing and many others followed his lead. He began what was later called the verbal learning tradition in the study of memory.

Another important tradition in the study of memory derives from a book by a British psychologist, Frederic Bartlett (1886–1969) in 1932. In *Remembering: A study in experimental and social psychology*, Bartlett performed rather casual experiments using prose materials and pictures. He argued that these were more like materials people encountered in their daily lives than were nonsense syllables and he made a point of testing people in more natural settings rather than under strict laboratory conditions. Bartlett showed how people would often recall material inaccurately, but that the errors were often plausible ones. For example, in recalling a native American folktale called ‘The war of the ghosts’ that contained supernatural details unfamiliar to his British college students, Bartlett noted that the students would often drop out these peculiar elements or, if the elements were recalled, subjects would give the elements a new meaning to make them fit better with the way they organized the story, a process he called rationalization. Bartlett developed the concept of schemata to explain the assimilation and later reconstruction of events. People interpret events of the world through their current knowledge structures, or schemata, and these organizational schemes also guide recall.

Bartlett emphasized the constructive nature of remembering: remembering was not the consultation of lifeless memory traces stored in the brain, but rather an active construction using bits of past experience woven together with current
expectations and knowledge. This emphasis on construction and studying more lifelike situations was different from the verbal learning approach and much later inspired researchers to look outside the laboratory to study the phenomena of human memory (Neisser, 1978).

Ebbinghaus and Bartlett were both pioneers, each developing unique styles of research and differing perspectives on remembering. Both have left a long legacy of research. Much of the best research in the field arises from a combination of these approaches, by researchers who attempt to take the interesting memory phenomena that can be observed in the world and to bring them into more tightly controlled laboratory situations for careful study.

### 7.2 Experiments on Memory: Logic, Methods, and Findings

Studies of human memory typically have at least two distinct phases: a study phase and a test phase. In the study phase people are exposed to some events or material and in the test phase they are queried in some way about their memory for the earlier experience. Occasionally, in naturalistic studies of remembering, people might be asked to recall events from earlier in their lives, such as the earliest memory that they can retrieve from their childhoods (Waldhof, 1948) or their classmates during their senior year in high school (Williams & Hollan, 1981). In the former example, the accuracy of the response cannot be checked (did the retrieved event actually occur early in their lives, or is it an experience described by a parent later?). In the second example, the recollections could be checked with yearbooks when available. However, in general, experimental psychologists like to gain control over the study phase of the experiment by presenting materials in the laboratory, like both Ebbinghaus and Bartlett did. When the experimenter has control over the conditions of study, then more secure inferences can be made about what is later retained (Schacter, 1995b).

#### The Encoding–Storage–Retrieval Framework

Psychologists have found it useful to conceive of the learning/memory process as composed of three stages: encoding (or acquisition) of information, storage (or retention), and retrieval (or accessing the memory). Encoding is the first stage that occurs during study and refers to accurate intake or perception of the event. Storage, or retention, is the second stage and refers to maintenance of the information over time; presumably, encoding an event produces some change in the nervous system and this change (often referred to as a memory trace) persists over time. Retrieval is the process of accessing the information later when it is needed. Whenever an event is successfully remembered, all three stages must be intact: the event was encoded, stored, and retrieved. When a person fails to remember an event, one or more of the stages failed. Pinpointing which stage failed is often difficult.

These three stages are logically necessary, but in practice it is difficult to separate them cleanly (Watkins, 1990). For example, when does encoding of an event end and storage or retention begin? How can we know exactly what information is stored after an experience? If an event cannot be remembered, how can we determine at which stage the fault lies? These questions cannot be answered conclusively. In addition, as we will see later in the chapter, the three stages are intertwined. The cues most effective for retrieval depend critically on how information was encoded and stored.

Although encoding and storage are difficult to separate, statements about storage and retrieval can sometimes be made based on experimental evidence. Consider an experiment by Tulving and Pearlstone (1966). They gave high school students lists of words to remember that belonged to common categories. So, students heard ‘Furniture – dresser, lamp; Birds – eagle, parakeet; Trees – elm, hickory,’ and so on, for 24 categories with two words in each category, so 48 words in all. The students were told to remember the words for a later test. After the list was presented, students were given a free recall test in which they were asked to recall words in any order; in this condition they recalled, on average, 18.8 words. So, information about this many items was successfully encoded, stored, and retrieved. We can then ask about the fate of the other 29. Were they never perceived or encoded in the first place? This seems unlikely, because the words were presented rather slowly, at the rate of one word per 2 s, but the possibility cannot be totally ruled out. Were the words encoded, but then somehow ‘lost’ from storage before the test? Or were the words (or some of them) successfully encoded and stored, but the subjects failed to retrieve them?

Tulving and Pearlstone (1966) employed a second condition to help answer the last question. They presented another group of subjects with the same lists under the same conditions, but tested them with a cued recall test rather than a free recall test. In the cued recall test subjects were given the category names (Furniture – ? Birds – ? Trees – ?) and asked to recall
list items in response to these cues. Now subjects recalled 35.8 items, or nearly twice as many as the free recall group was able to produce. The category names served as powerful retrieval cues, enabling subjects to recall much more information than they could without cues (free recall). But, you might think, perhaps the cues were not really effective but subjects simply guessed items in the category when given the category names. Tulving and Pearlstone were aware of this potential criticism and took several steps to rule it out. For example, the category members were not the first ones that occur on free association tests, but somewhat less common (and therefore harder to guess) members. So, for example, the birds were not robin and sparrow, but eagle and parakeet.

Based on these data, Tulving and Pearlstone (1966) made a critical distinction between information that is available in memory and that which is accessible. Any particular memory test (free recall, cued recall, recognition, and others, as discussed below) permits an assessment of the information that is accessible from memory, or the information retrievable under a particular set of test conditions. The information that is accessible, however, may not sufficiently represent the quantity of information that is available (or stored) in memory. No test provides an accurate assessment of exactly what has been stored in memory. We can say from the data presented above that in these experimental conditions more information was available for subjects tested with free recall than they were able to access. They free recalled only about 19 of the words but could have recalled around 36 under cued recall conditions. (In fact, the free recall subjects took a cued recall test after the free recall test and improved to about the level of the subjects given the cued recall test.) So, more information was available (stored) than was accessible (retrievable) under free recall conditions. Cued recall provided an estimate of 36 words recalled. But what about the remaining 12? Were they not encoded? Did they fade from storage? Or were they potentially retrievable with more powerful cues? We cannot say, short of finding the more powerful cues that might elicit their recall.

Distinguishing among encoding, storage, and retrieval is quite useful, even if there are gray areas in the use of these terms. Likewise, it is useful to distinguish between information that is available (stored) in memory and that which is accessible (retrievable) on any particular test, even though we can only measure the latter concept and not the former. At least the distinction prevents researchers from making the mistake, all too common in the history of memory research, of assuming that performance on a particular memory test provides an accurate measure of what a subject has encoded and stored. We will use these distinctions and terms throughout the remainder of the chapter.

Four Dimensions of Memory Experiments

Jenkins (1979) provided a useful way of thinking about experiments on human memory. He noted that researchers conducting experiments on the topic must make choices along four dimensions, whether or not that particular dimension is of interest in the experiment. These four dimensions are (1) the type of subject or person being tested, (2) the material used for the experiment, (3) the orienting tasks (or the features of the setting in which the people are tested), and (4) the type of test used to assess retention. If these four factors are placed in a plane and considered together, as in Figure 7.1, the appearance is one of a tetrahedron (although it is difficult to get this vision across on a two-dimensional page). The basic point is that all experiments on human memory involve choices in this four-dimensional space.

Consider the Tulving and Pearlstone (1966) experiment just described. The subjects were high school students, the materials were categorized word lists, and the subjects were all given

![Diagram](image_url)

**Figure 7.1** Jenkins' (1979) Tetrahedral Model of Memory Experiments. The variable of interest in a psychology experiment must be considered within the context of other variables which are held constant.
intentional learning instructions during the study or encoding phase, meaning that they were told to pay attention to the words because they would later be tested on their memory for the words. So, these three factors were all held constant. The fourth factor, the type of test used, was the independent variable of interest. Some subjects received a free recall test whereas others were given a cued recall test. Subjects tested with cued recall produced more words on the test than those tested with free recall.

Jenkins’ (1979) tetrahedral model of memory experiments illustrates that any experiment is created with some limited variables, or factors, of interest, and yet the study is conducted in the context of other variables, or factors, that are held constant. This position is referred to as contextualism: every experimental result is embedded in the context of other factors that were held constant. Appreciation of this point leads naturally to questions that should always be borne in mind. Would the finding from a particular experiment hold if other dimensions were manipulated? In the present case, if different types of subjects, or different materials, or other instructions prior to encoding were given, would cued recall still exceed free recall? There can be no a priori answers to such questions without the requisite program of research, yet researchers often implicitly assume that what they have found under their particular limited set of experimental conditions will generalize broadly to other conditions. Sometimes later research shows this to be true, but often later research will show that some exciting finding occurs only under a narrow set of conditions. The finding may still be of interest for theoretical or practical reasons, but its generality may not be great.

Jenkins’ (1979) model provides a useful way of surveying some main findings about memory. We will consider each of the four factors in turn and discuss some of the variables that have been manipulated. We will describe some of the variables that have been shown to have powerful effects on most standard tests of memory involving recall and recognition. However, our generalizations will come with a ceteris paribus clause (all other things being equal). That is, the generalization holds under most conditions, but it is often possible to manipulate factors in the other three dimensions that would at least affect the magnitude of the experimental effects. In some cases the effect could be eliminated or even reversed under other conditions. Because our net is so wide in this brief sketch, we cannot review the literature exhaustively. We consider materials, orienting tasks and the experimental setting, the type of subject, and finally the type of tests.

### Materials

Psychologists have studied memory for virtually every imaginable type of material. In the realm of verbal materials, lists of nonsense syllables and lists of words are common. In general, the more similar nonsense syllables are to real words, the better they are remembered. In addition, words that are frequently used are generally easier to recall than are words less frequently used. Phrases, sentences, paragraphs, and stories have been studied and, in general, the more meaningful and familiar the material is, the better it is remembered. Not surprisingly, when the words of a sentence are scrambled, they are remembered more poorly than when the sentence is presented intact. This points to an important feature of encoding: the ‘same’ material (the same words, in this case) can be well or poorly remembered depending on how it is presented. If it is presented in a coherent sentence it can be grasped and assimilated with past knowledge more easily. In the jargon of cognitive psychology, the sentence is more easily recoded (Miller, 1956). The same holds true even for lists of words. If they are organized in a meaningful fashion, they are more easily remembered than if presented randomly.

Retention for many types of nonverbal material has also been studied. Pictures, usually in the form of simple line drawings of common objects (a ladder, a helicopter) can be presented, with recall measured by having subjects produce names of the pictures. Pictures are better recalled and recognized than are the same items presented as words, even when the mode of response is verbal (which would seem to give the advantage to words, which are produced in the same form as the one in which they were presented). Faces are another popular material and some have argued that humans have special mechanisms for remembering faces, due to their importance to survival. In fact, some researchers have argued that the human brain has a special module or processor for faces and their retention, given their importance to human existence (see Bruce, 1988). Videotapes and complex pictures are sometimes used as stimuli for memory experiments, too. The idea behind using these types of stimuli is to study nonverbal memory; however, line drawings can be named, faces can be described (a bald man, with glasses, a high forehead) and so can complex pictures and videotapes. When psychologists want to study material where all reasonable types of verbalization are excluded, they turn to snowflakes (Neath & Knoedler, 1994) or kaleidoscope pictures (Wright, Santiago, Sands, Kendrick, & Cook, 1985) and use recognition procedures by testing with, for example, some snowflakes that
are the same as those that were studied and some that are different, with the task being for subjects to recognize the previously presented snowflakes.

Psychologists often present materials visually, but some types of material such as words or digits can be presented auditorily and often interesting differences occur between these modalities. In addition, various types of non-verbal sounds can be presented for later recall or recognition: music, tones, or various sounds (popcorn popping, birds chirping, etc.). In the olfactory modality, people can sniff various smells and be asked to recall or to recognize them later. In fact, olfactory memory is interesting because different patterns sometimes emerge than with other classes of stimuli. Rounding out the modalities, people can also be tested for various objects touched against their bodies or for various tastes, although these modalities are rarely used in memory experiments.

Orienting Tasks and Settings

This category of variables refers to many aspects of the experimental context, including instructions subjects are given, the particular version of a task they are provided, the strategies they might use, whether the independent variable is manipulated between subjects (so each subject receives only one condition) or within subjects (so each subject participates in each condition). The potential number of such settings variables is quite large.

In memory research, one basic instructional variable that must be decided for each experiment is whether to use intentional learning instructions (they are told they will be tested so they should learn the material) or incidental learning instructions (they are not forewarned about the later memory test when given the material). Suppose subjects are to be exposed to a list of 50 words and they are asked to rate the pleasantness of each word on a scale of 1 to 7, just to insure they pay attention to the words. Later, they will be asked to recall the words. If one group of subjects is warned, prior to the task of rating the words, that they need to remember the words for a later test, this would constitute an intentional learning condition. If the subjects were not forewarned, the condition would be one of incidental learning. That is, the learning of the list of words would be incidental to what subjects perceived their main task to be, viz., rating the pleasantness of the words. In general, intentional learning instructions produce greater retention than do incidental learning instructions, but this effect depends on the nature of the orienting task (rating pleasantness, in this example) given to the subjects. Actually, with pleasantness ratings, there would be little or no difference between incidental and intentional learning instructions. The reason is that when subjects rate words for pleasantness the task encourages thinking about the meaning of words, just as intentional learning instructions do. Because processing words for their meaning is a critical ingredient in remembering them, pleasantness rating encourages the type of processing that provides good retention and so, even with incidental learning instructions, recall and recognition of the words would be good (Craik & Lockhart, 1972). With tasks that do not encourage encoding of meaning, such as noting whether the word is printed in upper case or lower case letters, giving intentional learning instructions would typically reveal superior retention to incidental learning conditions.

Another factor of interest is the type of orienting task itself. Craik and Tulving (1975) had subjects exposed to words under incidental learning instructions. They were presented with a question and were asked to answer 'yes' or 'no' to the question in relation to the word that was presented immediately after the question. To take one example, consider the word BEAR given with the following three questions: Is it in upper case letters? Does it rhyme with CHAIR? Is it an animal? Subjects answered questions by pressing a key to indicate yes or no, as fast as possible. Other questions were similar to these (Is it in lower case letters? Does it rhyme with SOCK? Is it a type of bird?) but the answer was no. Later subjects were given a recognition test and required to select words they had recently studied from others that had not been presented. The results are shown in Figure 7.2 for items for which the answer to the question was yes. The outcome is dramatic: although all items were exposed for the same length of time, the split second encoding operation performed by the

Figure 7.2  Data from Craik and Tulving (1975). Those items encoded on a deep level were better remembered than those items that were encoded on a shallow level. Chance performance was 33%.
subjects in answering the questions greatly affected the later memorability for the words. Asking questions that required analysis of meaning (Is it an animal?) produced much higher recognition than did analysis of the word’s phonemic properties (Does it rhyme with CHAIR?), which in turn led to greater recognition than questions that called attention only to graphemic, visual properties of the word (Is it in upper case letters?). These results point again to the power of encoding processes: the outside world provides information to the senses, but whether it will be later remembered depends on the type of processing accorded the information.

Under intentional learning conditions, subjects can be given different strategies for use in learning material. Bower (1972) gave students lists of 20 pairs of words to remember, such as thumbback-pickle. One group was told to silently rehearse (to repeat over to themselves) each pair until the next pair appeared on the screen. A second group was told to form interactive images between the referents of the two words. That is, they should imagine a thumbback being pushed into a pickle. Later, both groups were given a test in which they were given the lefthand member of each pair (thumbback – ?) and asked to recall the righthand member. Bower found that people who simply repeated the two words recalled about 45% of the words whereas the group that formed images recalled close to 75%. So, the instruction to use mental imagery increased recall by 30% over the strategy that most people use for memorizing, rehearsing information to themselves! In general, recoding verbal information in terms of mental images is a powerful aid to memorizing and therefore this strategy is a prominent ingredient in many memory improvement books and programs. Rehearsing information often has small or negligible effects on recall, even when rehearsal occurs over prolonged periods.

Some memory improvement books suggest that forming bizarre images helps more than using ordinary images. In the above example, the thumbback and pickle should be made large and perhaps grotesque to engender the greatest mnemonic effect. However, the experimental evidence on this issue is mixed: some studies find that bizarre images help, whereas other studies do not. Why? One factor at work seems to be another contextual variable mentioned above, whether subjects are tested by having common and bizarre images in the same list or in different lists. When subjects form bizarre images to some items and common images to others, all within the same long list, they remember bizarre images better. However, when all items in the list are encoded either with bizarre or with common imagery, then often no difference exists in recall of the two types (McDaniel & Masson, 1985).

This outcome leads to a further point about encoding: when events are distinctive or salient, such as one bizarre image among many common ones, people often remember that event better. When events in our lives are powerful or emotional (either good or bad), we typically remember them very well. In fact, such events have been given the name flashbulb memories (Brown & Kulik, 1977) because they seem recorded as in a photographic flash. It is wrong to think of these powerful memories as literal photographs, because some research has shown that even when we think memories have this crystal clear quality, we can sometimes be quite wrong. However, distinctive, emotional events are usually remembered quite well.

Subjects

Psychologists have tested memory abilities of virtually all types of people. There are studies of college students, of children, of infants, of the mentally retarded, of depressed people, of schizophrenics, of older adults, of bilingual and trilingual individuals, of people with superior memories, of brain-damaged patients, of people receiving electroconvulsive therapy, and on and on. That said, we must admit that the greatest bulk of the research has been conducted with university students. Because most of the researchers in human memory are professors, students represent an abundant resource. We might make an analogy to the science of genetics by saying that college students are the Drosophila of human memory research. Just as geneticists use these fruit flies to work out the laws of genetics (fruit flies have relatively simple genetics, are easy to breed, and have short life spans), so researchers in human memory use college students. Like Drosophila to genetics researchers, college students have advantages for memory researchers as a population to study. They are usually intelligent, they are motivated, and millions of years of evolution have adapted them to their current niche in society: they participate in the educational system where they are given material to memorize for some 15–20 years, reaching expert level by the time they are in the university. Therefore, students represent an excellent model system for the study of human memory.

The problem of generality then arises: if the laws of memory are worked out with student populations, will they generalize to other groups? Of course, we might expect most other populations to perform worse over all than students, who are young, intelligent, and expert memorizers. In general, that is the case: young
children and older adults, for example, both show generally poorer recall and recognition than do college students. So do depressed people, schizophrenics, brain-damaged individuals, and so on. However, the critical question is whether the generalizations about the effects of variables change with these other factors. Considering some of the findings in the preceding sections, does the effect of forming mental images, or the advantage of meaningful processing, or the effect of sentential organization change dramatically when different populations are tested? The answer to all these questions is no. In fact, at the risk of too broad a generalization, we can say that the effect of most variables is similar across numerous subject populations. There are a few exceptions to this claim, one of which will be considered in the next section, but as a general rule, different populations of people respond to powerful memory variables in much the same way. If a variable (such as imagery) has a positive effect in one group (college students), it probably will also have a positive effect in another group (children).

One fascinating topic is whether there is a small group of people who have memory abilities spectacularly better than those in the normal population. Do some people have photographic memories? Although some people do have extraordinarily good memories, the word photographic is probably too strong. One remarkable mnemonicist was a Russian named Shereshevski, who was studied by Luria (1968). He could perform remarkable feats of memory and had great powers of mental imagery, as Luria documents in his fascinating book, *The Mind of a Mnemonist*. However, few others can duplicate his talents. Nonetheless, other mnemonists often have great powers, even if different from those of Shereshevski. An Indian named Rajan Mahadevan has a spectacular memory for numbers and has memorized pi (3.141592 . . . ) to over 31,000 decimal places! Since pi is nonrepeating, a test of Rajan’s memory is to give him five or ten digits and ask if this series of digits can be found in the first 31,000 digits of pi. He can decide yes or no within a matter of seconds with remarkable accuracy (Vogel & Thompson, 1995). Other mnemonists have specialized knowledge of the Bible or the Talmud, or have memorized telephone directories or dictionaries. Whether these remarkable people simply have radically different memory systems from the rest of us or instead use the same basic mechanisms as ordinary people but with greater efficiency is not known.

**Tests that Assess Memory**

Over the past century, many standard laboratory paradigms have been used to assess memory. We can broadly classify these tests as explicit memory tests or implicit memory tests. Explicit tests are, as the name implies, tests that provide direct or explicit queries about past experience. If subjects are asked to recall the list of pictures they recently saw in an experiment, their high school classmates, or the people they saw yesterday, these are explicit tests of memory. Implicit memory tests also measure the effects of past experience, but they do so indirectly through measuring transfer of past experience in current behavior without any instruction to retrieve from the past. For example, suppose you read the word *perspicacious* in a book yesterday. Today you are describing an intelligent acquaintance to one of your friends and you call her *perspicacious* in the course of the conversation. If you have never used that word before, you probably did it today because you read it yesterday. Therefore, we can say that reading it yesterday primed your use of the word today. Implicit memory tests are those that measure this kind of priming, although measured more formally than in this example. We first consider explicit measures of memory and then turn to implicit measures.

Explicit measures capture our conscious recollection of the past. Most of the experimental work on memory in the last century has used explicit measures. If people are given a list of 20 words and asked to recall them in any order, the task is free recall; if they were asked to recall them in the strict order in which they had been presented, the task would be serial recall. Both of these are explicit measures of retention. Paired associate learning was exemplified above; people are given pairs to study (*thumb-black-pickle*) and later given one member of the pair and asked to recall the other. In describing an experiment in an earlier section, we discussed another cued recall test in which subjects were given category names (such as Birds) as an aid to recalling words presented with them in a list (eagle, parakeet). Cues can also be used that were not included in the material to be studied. For example, if the word *elephant* had been presented in a list of words, *task* might be given as an extra-list retrieval cue for that item.

Recognition procedures have also been described in this chapter. In general, people are given material to study (say, 100 pictures) and then are tested by giving them the studied pictures and a large number of other pictures with the instruction to identify the ones they saw earlier. If they see the studied and nonstudied pictures one at a time and are instructed to say
yes or old and no or new, the test is called a free choice (or yes/no) recognition test. The subject makes a choice (yes/no or old/new) about each item. A forced choice recognition test involves giving subjects a studied item paired with one or more nonstudies items and asking them to pick the single studied item in the group. Free choice recognition tests are rather like true/false tests used in educational assessment, whereas forced choice tests are like multiple choice tests.

Throughout much of the history of research on human memory, people selected a convenient test and studied problems of memory via its use. Researchers could spend whole careers studying, say, processes involved in free recall or in recognition. The implicit assumption is that 'memory is memory' and that all tests reveal the same process, with some tests (say, recognition) more sensitive than others (say, free recall). Test paradigms were considered as what other scientists called preparations: once one has developed a sensitive task or preparation, it is used repeatedly. Although this research strategy can produce dividends in deep understanding of a limited range of problems, in the long run it seems misguided. The reason is that tests are not equivalent and do not measure the same 'thing' (such as strength of a memory trace). Often, the same variable can be shown to have different effects on various measures of memory. Consider the effects of frequency of words in the language. As noted above, high-frequency words are better recalled than low-frequency words when other characteristics (part of speech, length, etc.) are held constant. We might be tempted to explain this by saying that high-frequency words create stronger memory traces than do low-frequency words. However, when recognition measures are used, it turns out that low-frequency words are better recognized than are high-frequency words! For other variables there might be an effect on recall, but no effect on recognition. Therefore, recall and recognition cannot be considered as measuring 'the same memory' with differing sensitivity.

A useful perspective for thinking about the relation between the encoding and testing of memories is transfer appropriate processing (Bransford, Franks, Morris, & Stein, 1979). In general, performance will be better on tests to the extent that the process engaged by the test match those used in encoding the materials. Let us consider an example. In an experiment by Stein (1978), subjects studied words such as 'knife' and answered one of two questions about it. In one case, the question asked about visual features ('Does the word have a capital i in it?') whereas the other type of question forced subjects to think about the meaning of the word, such as in 'Does ___ have a steel blade?' This experiment is similar to one we discussed earlier in which we reported that, on a standard recognition test, questions encouraging the processing of meaning led to better recognition memory. However, Stein's experiment used two different kinds of recognition tests. One was the standard forced choice test in which subjects picked out a studied word from among nonstudied words (e.g., trUck, knIfe, relAy, and sCene). As can be seen in Figure 7.3 on the left, the prior work was replicated on this test, with semantic encoding conditions producing better recognition than the condition that required people to examine letters. However, consider the other kind of recognition test in which people were tested on their ability to recognize the visual appearance of the words they had just seen. Now the test consisted of a forced choice test like this – kNIfE, knIFe, knIFe, knIFe – with subjects required to pick the form of the word they had studied. As shown on the right side of Figure 7.3, performance on this recognition test was the opposite to that on the standard test, with better recognition following prior attention being drawn to the appearance of the word during the encoding phase. Despite the fact that subjects encoded the same words in the same ways, the tests brought out different aspects of the encoded experience. A test that measured recognition of visual appearance was aided more by prior attention to the visual appearance of the words than by attention to their meaning; a test that measured recognition of meaningful aspects of prior experience benefited more from prior attention to the meaning of the words than by attention to their appearance. This pattern of results exemplifies the principle of transfer appropriate processing.

As noted above, implicit memory tests measure retention indirectly. The study phase of the typical form of these tests is much like it is for
explicit memory tests. Usually subjects are exposed to either words or pictures under either intentional or incidental learning conditions. However, the test is quite different from those of explicit tests, because subjects are not told to remember events from their past. Instead, they are given another task to perform that is ostensibly unrelated to prior events in the experiment. For example, if subjects have been exposed to a list of words such as elephant, accordion, and thimble, then sometime later in the experimental session they might be told that their next task is to complete fragmented words with the first word that comes to mind. One version of the task, called word stem completion, provides subjects with the first three letters of words and asks them to say the first word that comes to mind. Some of the stems would correspond to previously studied words (ele____, acc____, th____) and many other stems would not. In a different version of the task, called word fragment completion, subjects are provided with fragmented forms of words such as e_c_h_n_ and asked to guess what the words are, again with some fragments referring to previously studied words and others not. In both these tasks, the results of hundreds of experiments demonstrate that subjects are more likely to complete the word stems or word fragments with words that have been recently studied. For example, if elephant has not been studied, the probability of completing the word fragment with this word might be .30, whereas if the word had been studied, the completion rate might be .60. This 30% benefit is attributed to priming; prior study of the word primed word completion.

Psychologists became interested in these implicit memory tests because the outcome of experiments using implicit memory measures is often quite different from those using explicit measures. Implicit measures of memory reveal the influence of prior experience on behavior and so reflect a type of memory, but it is a very different type from that revealed by explicit memory tests such as recall and recognition. For example, some patients who have suffered brain damage lose most of their ability to remember new information, a condition called anterograde amnesia. Their forgetfulness is extreme; if they receive a list of words to remember and then are asked to recall or to recognize the words after a short retention interval, they may not even remember having heard a list, much less the particular words that were in the list! Performance on all explicit memory tests is quite poor for such patients.

The natural interpretation of this amnesia is that the centers in the brain responsible for encoding and storing new information have been damaged. However, in the 1960s British psychologists Elizabeth Warrington and Lawrence Weiskrantz (1968) discovered that such patients showed perfectly normal patterns of priming on implicit memory tests, a finding replicated many times now. The results of one such replication by Graf, Squire, and Mandler (1984) are shown in Figure 7.4. Subjects in the experiment were either amnesic patients or control subjects matched on age and level of education. All subjects were shown lists of words and then given either a free recall test or a word stem completion test. As seen on the left side of Figure 7.4, patients were much worse than controls on the free recall test, which is no surprise. However, in the word stem completion results shown on the right, patients showed as much priming as did control subjects. The benefit from prior study of the words was normal for these patients, which may indicate that encoding and storage of information was intact. The amnesic patients’ difficulty seems to be in retrieving the information when asked to remember it; the processes of conscious recollection have been disrupted. This pattern of results reveals a dissociation between measures of memory: performance on one set of tasks (explicit memory tasks) is greatly affected by brain damage, whereas performance on the other type of task (priming on implicit memory tests) is affected very little, if at all. This outcome represents an exception to the generalization made in the previous section that most individual difference variables affect the overall level of performance in memory experiments, but do not interact with these variables.

Many experimental variables also affect explicit and implicit measures of memory differently. For example, manipulation of intentional versus incidental learning instructions during study has little effect on implicit tests such as word stem and word fragment completion and
neither does study time for individual items, massed presentation of items, and other factors such as the type of orienting task (rating words for appearance or for their meaning). The study of priming on implicit memory tests has burgeoned over the past 20 years and there is now a substantial literature on the topic, with general references provided in the Resource References. We should note here that there are many different types of implicit memory tests besides word stem and word fragment completion, so the claims we made above hold for these tests but not necessarily for all implicit tests.

We close this section by returning to the theme at the beginning: we cannot talk about ‘memory’ as if it is one entity. There are many ways of measuring memory and they reveal different processes and systems at work. It is misguided to think of memory as being all of one type, or of variables having a general effect across all types of settings, tasks, and people. Interactions or dissociations among measures of memory are the rule and not the exception.

7.3 Theories of Memory

There are dozens of theories of memory, which represents something of an embarrassment of riches. The presence of so many theories in the field reflects little agreement on which theory (or even which approach) is correct. Therefore, our aim in this section is to describe several categories of theories of memory rather than just one or two theories. First, however, we consider a theory that is widely accepted and is built into our way of talking about memory, but which is known to be untrue.

Strength Theory

One simple and appealing idea, built into many theories, goes like this: experiences in the world are coded into the nervous system in terms of memory traces, which vary in strength; some traces are strong (leading to good retention) whereas others are weak. For example, in a simple memory experiment in which subjects are presented with 30 words for later recall or recognition, some words might be said to leave stronger traces than others. These would be the words that tend, across subjects, to be well recalled and well recognized. Other words would lay down weaker traces. A representation of this idea is shown in Figure 7.5. A word must have considerable trace strength to be recalled because recall requires more strength than does recognition. The abscissa in the figure represents hypothetical traces of words labeled A, B, C, D, and E whereas the ordinate represents strength of these traces. The figure accounts for why some words are recalled and others are not; similarly, the effect of experimental variables that generally increase retention, such as study time, can also be described as increasing trace strength. Words studied longer have stronger traces than those studied for less time.

The theory as portrayed in the figure can also account for the finding that often words can be recognized if they cannot be recalled. The threshold (represented by the horizontal line) for the amount of strength needed to succeed on a recognition test is lower than that for a recall test. More trace strength is required for recall than for recognition.

The simple strength theory of memory, as described in the previous paragraphs and as portrayed in Figure 7.5, can explain many basic facts about memory: why some items are recalled and others are not, why experimental variables affect retention, and why recognition is often more sensitive than recall, in the sense that items can be recognized when they cannot be recalled. The theory is simple and accords with our common sense way of describing memories as strong or weak. So, what is wrong with it? Basically, although it accounts for some facts, the theory fails to explain most of the interactions or dissociations described in the preceding sections. If the sole determinant of retention is memory strength, how does one account for differing patterns of performance as a function of the type of test given? According to the simple strength theory, tests differ only in their threshold for performance. In an earlier section, we noted that high-frequency words are better recalled (have higher strength) than low-frequency words. On the other hand, low-frequency words are better recognized than
high-frequency words. Strength theory cannot explain such effects very easily, without making assumptions that undercut the theory. How can one and the same trace be stronger on a recall test but weaker on a recognition test? Similar types of problems crop up in comparing priming on implicit memory tests with recall and recognition on explicit tests, because such comparisons often reveal very different patterns of performance on the two types of test. There is no simple way to account for these differences in terms of a strength account and we can therefore reject it (for these and other reasons). However, the idea of memory strength is remarkably resilient and survives in many modern theories despite the fact that it has never worked. We consider below some promising alternatives, ones that try more realistically to deal with the variety and complexity of the data about memory.

**Interference Theory**

Interference theory was developed to explain forgetting. Strength theory’s simple (and vacuous) account of forgetting is that, over time, the strength of traces simply wanes and therefore forgetting occurs. McGeoch (1932) mounted an attack on this simple decay or disuse theory of forgetting and argued that interference theory provides a more potent explanation.

There are two main types of interference, both quite important in certain situations. Proactive interference refers to the inhibiting effects that prior events have on recall or recognition of events learned later. Retroactive interference refers to the inhibiting effects on memory for some target event of other events occurring after the target event. So, for example, suppose you park your car in a large lot every day and your task each evening is to find it. The fact that you will sometimes go to where you parked it yesterday reveals proactive interference, or the interfering effect of prior events. If you were asked to determine where you parked your car seven days ago, retroactive interference would also come into play: the places you had parked the car in the intervening six days would also interfere with your recollection.

Proactive and retroactive interference are descriptive terms, referring to general classes of events that affect retention. The mechanisms believed to be responsible for interference are response competition (for both proactive and retroactive interference). That is, when asking “where did I park my car a week ago?” the prior events and more recent events all compete as answers and the problem is selecting the correct answer from among these competitors. Another factor, unlearning/recovery, also has been invoked to explain retroactive interference. The general idea is that learning a new experience (parking your car today) may cause forgetting (“unlearning”) of previous experiences, such as where you parked it yesterday, which causes forgetting but also helps to update memory. However, unlearning does not seem permanent, because these inhibited responses can be shown to increase over time (Wheeler, 1995) and therefore to exert interference. Anderson and Neely (1996) provide a recent overview of inhibitory factors in remembering.

**Memory Systems Theories**

The guiding assumption of this class of theories is that several different cognitive systems underlie human memory. It is wrong to think of ‘memory’ as a single entity; rather, there are multiple memory systems and the reason that performance can differ so dramatically across various measures of memory is that different systems are responsible for these effects. Among the most generally agreed upon memory systems are the following: short-term memory (sometimes called primary memory or working memory), episodic memory, semantic memory, procedural memory, and perceptual memory. Let us discuss each in turn, albeit briefly.

Short-term or working memory, as the name implies, is responsible for the ability to hold, maintain, and use information over short periods of time. For example, the span of immediate memory for most people is around seven items, plus or minus two (Miller, 1956). That is, for stimuli such as words, digits, or letters, most people can remember between five and nine items in the correct order, with the average being seven (which is why local phone numbers are seven digits long). This short-term (or working) memory capacity seems to be different from other types of memory in the sense that working memory has been shown to remain intact when other types of memory are impaired. For example, amnesic patients (who have an obvious memory impairment) are often able to remember telephone numbers for a short time and other types of information demanding use of working memory. There may be more than one type of working memory, specialized verbal and nonverbal systems. Baddeley’s (1986) influential theory of working memory postulates these distinct capacities and offers a systematic view of the available evidence.

Episodic memory (sometimes called long-term memory for events) refers to a person’s ability to recollect the events of one’s life. The hallmarks are that the events are personal ones
and that they occurred at a particular time and a particular place in the past. Therefore, to recollect these events, an individual must be directed to (and able to) retrieve from a particular time and place. When subjects are asked to remember the words from a list presented five minutes ago, or to recall their first week of college, or to remember their first trip on an airplane, these are all queries of the episodic memory system. It is this system that is damaged in amnesic patients. They can remember events over the short term, but they cannot retrieve from long-term episodic memory.

Semantic memory refers to general knowledge of the meaning of facts, people and events. If you are asked for the definition of a platypus, or who Winston Churchill was, or the capital of Japan, these would all be queries of the semantic memory system. When new episodes occur in our lives, we interpret them through our past knowledge (through semantic memory) to represent them as meaningful episodes, but semantic memory and episodic memory seem otherwise separate. To return to amnesic patients, often their brain damage does not affect semantic memory – they still know the meaning of words, history they learned in school, and they might still be able to play chess well if they could do so before their injury. The problem they seem to have is in remembering personal events. Semantic memory is ahistorical and impersonal; recalling the time and place of the occasion when you first heard about Winston Churchill would not be necessary to knowing who he was and what he did.

Episodic and semantic memory are together sometimes called declarative memory. These systems permit people to make declarations about their past. We say ‘I remember that . . .’ when referring to episodic memories, or ‘I know that . . .’ when referring to semantic memories. Non-declarative memory systems are all the other systems.

Procedural memory is considered the system responsible for motor actions. Riding a bicycle, serving a tennis ball or, for that matter, walking and running are complex motor skills that must be learned. The systems that support these processes are different from others we have considered above.

Finally, there are also perceptual representation systems that maintain a memory for the perceptual objects of the world as they impinge on the individual. These representations are believed to be stored in brain areas in which the experiences were first processed in the visual, auditory, or other perceptual systems. For example, priming in implicit memory tests such as word fragment completion and word stem completion is mediated by such perceptual representation systems.

These five systems do not exhaust the possibilities, which may extend to as many as 15–20 in some counts. However, the ones listed above are generally agreed upon. We next consider an alternative approach.

Processing theories

In processing or procedural approaches, the basic assumption is that the brain/mind is composed of many different types of processes or procedures that can function in many ways. What we call ‘memory’ is a very general property of cognition in that experience changes the way a person interacts with the environment. The processing approaches try to account for the many different kinds of mnemonic skill by invoking the idea that any task has many component processes involved in it. Each task has a somewhat different set and arrangement of components from any other task.

The components of a task are difficult to define precisely, except at a broad level. For example, some memory tasks seem to respond strongly to manipulations of meaning or to conceptual information. These tasks (like recognition memory, for example) are not much affected by changes in perceptual/surface characteristics such as modality of presentation (auditory or visual) of information, but they are greatly affected by differences in orienting tasks (judging the case of words versus a meaningful judgment such as rating pleasantness). Therefore, recognition seems to represent a conceptual or meaningful task. On the other hand, implicit memory tests such as word stem and word fragment completion are affected in an opposite way by these variables. Manipulations of meaning have little effect, whereas manipulations of perceptual characteristics of the stimuli, such as modality of presentation, have a large effect on these tests. Therefore, these implicit memory tests are called perceptual tests, whereas explicit tasks such as recognition and free recall are referred to as conceptual tests (Blaxton, 1989). Perceptual processes are preeminent in the former tests, conceptual processes in the latter tests.

A general point made by processing theorists is that there are no dedicated memory systems in the human brain. Rather, there are a variety of all-purpose systems, with one purpose being to retain after-effects of experience. The perceptual memory systems, from this perspective, are perceptual systems used to analyze incoming information from the outside world and that also
happen to change as a function of past experience. Procedural memory can be viewed in the same way. The nervous system changes as we practice a skill and these changes can be seen as procedural memory. However, all systems use procedures. Even in episodic memory experiments, how the query is presented to retrieve past experience determines the procedures used and the answer provided.

The emphasis on processing approaches to memory is the interactive nature of cognition. Although it is useful for some purposes to dissect the cognitive world into different capacities (perceiving, attending, remembering, thinking), all these processes interact. We cannot think about a problem without perceiving its features and having them call up ideas from the past. Similarly, we cannot perceive a complex scene without considerable past experience or knowledge that is stored in the brain. Perceiving requires memory just as memory requires perceiving. The processing approaches to cognition, sketched here, try to confront the interactive complexity of cognition (see Resource References for references on these points of view).

**Mathematical Models**

Theorists who endorse the memory systems and processing points of view attempt to cover wide territory, to account for differences from many tasks, many subject populations, and many variables. Theorists who use mathematical models typically endorse a different strategy as fruitful. This approach often considers rather limited situations, such as recognition memory, but develops highly specified models for the task of interest. The assumption is that proceeding this way – providing specific models that should be testable – will provide a useful way of conceptualizing memory. Thus, for example, exquisite and detailed models have been developed for the Sternberg (or item-Recognition) paradigm in which a series of digits is given (6 1 8 5 3 9) and then after the list presentation, one digit is given as a test item. If the digit matches (8), the subject pushes a yes key as quickly as possible; if it does not match (4), the no key is pressed. Much interesting work has come from studying this item recognition task (Sternberg, 1975) and elegant models have been proposed (and falsified).

More ambitious models have attempted to account for a range of phenomena. One such model is the Search of Associative Memory (SAM) model by Raaijmakers and Shiffrin (1981). They proposed that memory could be conceived as a large associative network and further specified a number of assumptions about how associations are formed during study episodes and retrieved during recall and recognition tests. The model has been used to make quantitative predictions in many different kinds of recall and recognition experiments. Although we cannot review the achievements here, the theory has impressively predicted a large range of results and shows the power of mathematical modeling in furthering our understanding of memory.

In this section of the chapter, we have provided several general approaches to theory development within modern cognitive psychology. As we have noted, many theories of memory exist and theorists approach the topics from different perspectives and select various patterns of results to explain. There is no single comprehensive theory of memory that has met with general acceptance.

### 7.4 A Sample Experiment: Roediger and McDermott (1995)

Psychologists interested in memory have usually measured accuracy of recollections in recall and recognition. The primary error they have examined is the error of omission, or forgetting – the inability to recollect something once known. However, another class of errors that people make is that of commission rather than omission – people can remember events differently from the way they happened or, in the most profound case of error, they can remember events that never happened at all. Thus far this chapter has dwelt on processes affecting accuracy in performance. Therefore, for the sample experiment we will focus on an experiment in which subjects make frequent errors.

Roediger and McDermott (1995, Experiment 2) tested 30 undergraduate students in a straightforward paradigm using 24 lists of words. The lists were produced from word association norms and the words were strongly associated. For example, one list was smooth, bumpy, road, tough, sandpaper, jagged, ready, coarse, uneven, riders, rugged, sand, boards, ground, and gravel. These words are the ones produced in a word association test to the word rough, but notice that rough is not included in the list. We refer to rough as the critical nonpresented item. In the experiment the 24 lists were arbitrarily divided into three sets of eight and each subject studied 16 lists. The lists were presented auditorily at the rate of one word every 1.5 s, after which the subjects received a signal that indicated they should either recall the list for 2 min or perform mental arithmetic for 2 min. After
Figure 7.6 Data from Roediger and McDermott (1995), Experiment 2. The probability of recalling the critical, nonpresented lure was slightly higher than the probability of recalling one of the words presented in serial positions in the middle of the list.

eight lists subjects recalled the words in any order as soon as the list had been presented (free recall), whereas after the other eight lists they performed arithmetic problems. The reason for doing this was to examine the effects of recall on a later recognition test. Before describing the recognition test, we describe the results obtained in recall.

Presented in Figure 7.6 are the primary results from the recall test. The function shown is called a serial position curve, because it plots probability of recall on the ordinate and the ordinal position of the presented item on the abscissa. So, the first point is the average probability of recall of the first item in the lists, and so on. Note that recall of the first few items and the last few items is better than that for the middle items. This characteristic almost always occurs in immediate retention of a series; good recall of the first few items is called a primacy effect and superior recall of late items is a recency effect. However, for present purposes the dashed line represents the data point of most interest. This line represents false recall, or recall of the critical nonpresented word (rough in our sample list above) that was not presented in the list. Obviously, average recall of this item was very high, at 55%. In fact, false recall of the critical nonpresented words was actually higher than for words that were presented in the middle of the list! Later experiments have generally shown that free recall of the critical nonpresented word is about the same as, or a little higher than, recall of the words presented in the middle of the list.

After hearing 16 lists and recalling eight of them, the subjects then received a recognition test in which they were instructed to decide whether each word was old (studied in the lists) or new (not studied). In addition, for each word they said was old, they were asked to judge whether they could remember when the word originally occurred in the list or if they just knew that the word had occurred but could not actually remember the moment of its occurrence. This remember/know judgment captures a person’s experience during recollection and is often used to measure remembering of experiences as distinct from the feeling of familiarity that can occur without true remembering (Tulving, 1985). The recognition test consisted of 96 words, 48 of which had been studied (targets) and 48 of which had not been studied (lures or distracters). Among the lures were the 16 words that were the critical lures, words such as rough that were associated to list words but not actually presented.

The recognition results are presented in Table 7.1. The top two rows represent correct recognition (calling a studied word ‘old’, or the hit rate) whereas the bottom three rows represent false recognition (that is, calling a word that was objectively nonstudied ‘old’; this is called a false alarm). Looking at the first column, which shows overall recognition, the most striking result can be seen by comparing the data in the top two rows, representing correct recognition

<table>
<thead>
<tr>
<th></th>
<th>Recognition</th>
<th>Remember</th>
<th>Know</th>
</tr>
</thead>
<tbody>
<tr>
<td>Targets</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study + recall</td>
<td>.79</td>
<td>.57</td>
<td>.22</td>
</tr>
<tr>
<td>Study + arithmetic</td>
<td>.65</td>
<td>.41</td>
<td>.24</td>
</tr>
<tr>
<td>Critical lures</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study + recall</td>
<td>.81</td>
<td>.58</td>
<td>.23</td>
</tr>
<tr>
<td>Study + arithmetic</td>
<td>.72</td>
<td>.38</td>
<td>.34</td>
</tr>
<tr>
<td>Unrelated lures</td>
<td>.14</td>
<td>.03</td>
<td>.11</td>
</tr>
</tbody>
</table>
of targets, with that in the next two rows showing false recognition of critical lures. (The first and third rows represent the case in which recognition occurred after recall; the second and fourth rows indicate recognition that did not follow recall.) Remarkably, the false alarm rate for the critical lures equaled or even slightly exceeded the hit rate for the target items. In addition, prior recall of the studied list increased both the hit rate and the false alarm rate for overall recognition. Both the hit rate for targets and the false alarm rate for critical lures greatly exceeded the false alarm rate for lures from lists that were not studied, whose data are shown in the bottom row.

The second and third columns in Table 7.1 shows the breakdown of the overall recognition rate into remember and know responses. Look first at the data for unrelated lures in the bottom row and see that the false alarms are usually deemed to be items known to be on the list, but that their actual occurrence cannot be remembered. This pattern makes sense, of course, because the lure items were not actually on the list. However, the top four rows show a very different pattern. For the target items, subjects say that they can remember the moment of occurrence for most items that were called ‘old’. That makes sense, because the items actually had been presented. However, exactly the same pattern occurred for the critical lures: despite the fact that they had not been studied, subjects not only falsely recognized them as having been presented, but also said they could remember the actual occurrence of the words on the list! In addition, when subjects had previously recalled the list, the illusion of remembering was even stronger.

The Roediger and McDermott (1995) experiment shows a simple laboratory technique that produces dramatic levels of false recall and false recognition. Unlike other work showing distortions of memory, this experiment was carried out under conditions that should encourage accurate responding. Subjects were tested for recall immediately after list presentation with an instruction warning them not to guess and the technique used (free recall) is often thought to be relatively error free because subjects set a high criterion for responding (and indeed few errors occurred, except for recall of the critical nonpresented word). In addition, the recognition test occurred relatively soon after the recall test. The instruction for subjects to make remember/know judgments is believed to make subjects more aware of the basis of their responses and to make them more carefully monitor their responses for accuracy. In addition, use of word lists (rather than prose or slide sequences) is often thought to encourage accuracy. Nonetheless, despite all these conditions that normally encourage accuracy, high rates of false recall and false recognition occurred.

The most likely interpretation of the Roediger–McDermott (1995) results is that subjects thought of the critical word during presentation of the list and then later confused their private thoughts with the overt presentation of list words. The general point is that people are always drawing inferences, going beyond the literal information in their environments. What we remember of experiences may be some blend of what actually happened and what we were thinking while the events were happening.

### 7.5 Applications of Research

Research on human memory has been applied in many settings. Because the coverage in this chapter has been selective, not all the basic research that could be applied has been covered here. We will provide vignettes of applications to issues of practical importance. Because the issues of memory pervade thought in many different contexts, the implications of basic research for these applied topics can be wide-ranging.

The demands on memory in modern life are greatly increasing. Complicated telephone systems, computer systems, fax systems, cameras, videorecorders, automobiles and the like give rise to increasing cognitive demands, often on memory. Simply the number of passwords and telephone numbers that must be retained exceeds the capacity of many people. Psychologists’ expertise is often brought to bear on these practical problems. In particular, telephone companies and computer companies hire many psychologists to work on making equipment useable by taking the human factor into account. These human factors psychologists, as they are called, try to engineer complex systems so that they can be used with relatively few errors by their human operators. For example, as discussed above, early research sponsored by telephone companies in England and the United States showed that the largest number of digits that could be remembered without error on an immediate test was seven, so telephone numbers were made seven digits long. Similarly, these psychologists showed that numbers are better remembered if they are grouped or chunked, so the typical way telephone numbers in the United States are presented is in groups of three, then four, as in 792-3948. In France, telephone numbers consist of eight digits and are presented as four sets of two-digit numbers: 79-23-94-81. Human factors psychologists must consider
other processes besides memory, but complex mechanisms created by engineers must be configured with attention to the limited capacity of working memory of the humans who use them. Psychologists often help devise the person/machine interface.

Research on human memory is also applied by educational psychologists. After all, one central issue in education (perhaps the central issue) is transmission of knowledge from the instructor and textbook to the student, so that the student can learn and remember the information. Therefore psychologists have spent much time studying how to organize and present information in textbooks, how to lecture, and similar topics. One problem is transmitting information obtained from basic research on human memory to educators who might be able to use it. Still, many findings from psychology have found their way into educational practice.

Psychologists have also designed programs of cognitive rehabilitation. For example, many medical conditions, such as Alzheimer's disease, impair cognitive functioning in general and memory functioning in particular. Cognitive psychologists have devised programs that will ameliorate (but certainly not cure) severe difficulties that people have in remembering.

Relatedly, psychologists have also studied memory improvement techniques and have developed courses and written books on memory improvement that offer general advice. Many mnemonics have been known since ancient times, but modern cognitive psychologists have studied specific aspects of advice provided in these systems to see which features of the techniques really work. For example, if imagery is used in a mnemonic device, is it necessary for the images to be bizarre ones for them to be effective? As discussed above, the answer is not clear-cut, but selective use of bizarre imagery on especially difficult items should improve performance.

Cognitive psychologists who study memory are also often called upon as expert witnesses in legal situations that turn on the vicissitudes of human memory. Three types of case will be mentioned here. First, many cases turn on eyewitness identification. If a witness to a crime identifies a suspect as the culprit, this fact represents powerful evidence in the courtroom. But how accurate are eyewitnesses? Much research has been conducted on this topic, with the finding that witnesses can sometimes be wrong, but quite confident that they are right. Conditions that foster mistaken eyewitness identification have been discovered and psychologists testify about this research in court cases.

A second type of court case concerns children's testimony. Can children serve as reliable witnesses? How accurate are their recollections? Under what conditions are they likely to be accurate? When might they be misled into making false accusations? These issues are central in cases in which children are the primary witness or the only witness. Therefore, developmental psychologists who study memory development serve as experts in these cases.

A third type of court case, much in the news in the United States, Canada, and the United Kingdom in recent years, concerns cases of recovered memory after long delays. In a typical case a woman in her 30s or 40s undergoes psychotherapy for depression or some other problem. The therapist may inquire about her experiences during childhood in the normal course of therapy. In the process of talking about her childhood, or in response to questions and demands from the therapists, the patient may recover memories of unpleasant events from that period. In the most dramatic cases, she may remember being physically or sexually abused by her father, by an uncle, or by a teacher. In some cases, these recovered memories during therapy are the first inkling that the patient had of such abuse. One interpretation is that the events experienced were so horrible that they were repressed, or banished to an unconscious state, and then recovered during the course of therapy. Another interpretation is that the events never happened at all and the 'memories' were manufactured during the course of therapy, perhaps via suggestions from the therapist and other mechanisms (Lindsay & Read, 1994).

The issues in such recovered memory cases are very complex and present challenges to the legal system when they are taken to court. Psychologists testify as expert witnesses. Although different shades of opinion exist on the general veracity of these recovered memory claims, many researchers believe that recovered memories of abuse should only be accepted as critical testimony when there is converging evidence of abuse from some other source (hospital records, other witnesses, etc.). The vagaries of memory are well known, especially after long delays, and much evidence points to the suggestibility of memory as a potent source for these memories.

7.6 Conclusion

This chapter has provided an overview of the experimental psychology of human memory. As noted at the outset, research in several other traditions also contributes important insights to our understanding of human memory. These
include studies from neuropsychology (study of brain-damaged patients), from neurobiology (studying neural underpinnings of memory, usually with animal models), from neuroimaging (studying neural processes in intact humans whose brains are scanned by modern neuroimaging techniques while they are learning or retrieving information), from developmental psychology (studying memories of children and older adults), and from other directions, too. The study of human memory has become increasingly interdisciplinary. Keep in mind, then, that the perspective offered in this chapter is only one of many that are useful. However, the insights and research of experimental psychologists have defined the study of human memory in important ways.

RESOURCE REFERENCES

Bjork, E. L., & Bjork, R. A. (Eds.). (1996). Memory. San Diego, CA: Academic Press. This is an excellent collection of chapters on virtually all the topics included in this chapter and many more.


ADDITIONAL LITERATURE CITED


Memory Processes


Communications concerning this chapter should be addressed to: Professor Henry L. Roediger, Department of Psychology, Washington University, One Brookings Drive, St. Louis, MO 63130-4899, USA