HUMAN MEMORY AND COGNITION
To the three who matter most:

Mary Helen Homerin Ashcraft,
Jordan Mark Ashcraft, and
Laura Catherine Ashcraft.

Credit lines for copyrighted materials appearing in this work are placed in the "Credit" section at the end of this book.

Copyright © 1989 Scott, Foresman and Company
All Rights Reserved.
Printed in the United States of America.
23456 KPF 89089

Library of Congress Cataloging-in-Publication Data

Ashcraft, Mark H.
Human memory and cognition.
Includes bibliographies and index.
1. Memory. 2. Cognition. I. Title.
BF737.A68 1988 153 88-15838
ISBN 0-673-18862-0
Chapter 5

EPISODIC LONG-TERM MEMORY

• Preliminary Issues
  Metamemory
  Mnemonic Devices

• The Ebbinghaus Tradition of Memory Research
  Evaluating the Ebbinghaus Tradition

• Storage of Information in Episodic Memory
  Rehearsal
  Two Kinds of Rehearsal
  Depth of Processing
  Challenges to “Depth of Processing”
  Organization in Storage
  Imagery
  Storage Summary—Encoding Specificity

• Retrieval of Episodic Information
  Decay
  Interference
  Retrieval Failure
  Retrieval Cues and Encoding Specificity

• Autobiographical Memories
  The Relationship of Laboratory to “Real-World Memory”
Memory is the most important function of the brain; without it life would be a blank. Our knowledge is all based on memory. Every thought, every action, our very conception of personal identity, is based on memory... Without memory, all experience would be useless. (Eridge-Green, 1900)

We must never underestimate one of the most obvious reasons for forgetting, namely, that the information was never stored in memory in the first place. (Loftus, 1980, p. 74)

The fact is that we have almost no systematic knowledge about memory as it occurs in the course of ordinary life... Until we know more about memory in the natural contexts where it develops and is normally used, theorizing is premature. (Neisser, 1976, pp. 141-142)

A couple of years ago, I decided to go ahead and start using the special card my bank had sent me for its automated teller machines. It had been at least two years since the bank started issuing these cards, so I had completely forgotten the personal code number I had supplied to them for my account. After the bank sent me a copy of my code number, I started wondering how I had originally come up with that particular code—it wasn't an obvious number like my birthdate or address, and yet I doubted that I had just chosen the digits out of thin air. Finally, I pieced together what had happened. I had been teaching a special course on memory techniques, mnemonic devices and so forth. One of the standard mnemonics I had covered was the digit-to-letter conversion system, where numbers are translated first to letters, then the letters are used as short-hand words; for instance, 9 translates to t, 2 translates to n, so the number 92 could be remembered by the word “pan.” Once I realized that this mnemonic might be the source of my bank code number, I went back to the now-forgotten conversion table to see what word I must have had in mind when I chose that code number. As it happens, my bank code number translates back into a word that I should have been able to remember—“mnemonic.”

This is the first of two chapters specifically devoted to long-term memory, the relatively permanent storage vault for a lifetime's worth of knowledge and experience. Why do we need two separate chapters? Because there are at least two broad classes of knowledge that humans store in long-term memory, at least two kinds of long-term memory information. Likewise, there are two corresponding kinds of long-term memory research that cognitive psychology has performed. This two-part classification of knowledge was explicitly discussed by Tulving in 1972, in an influential chapter entitled “Episodic and Semantic Memory” (although as Tulving pointed out in his 1983 book, it was a distinction that was implicitly acknowledged and agreed upon well before his 1972 chapter).

In Tulving's classification, episodic memory refers to a person's autobiographical memory, to the personally experienced and remembered events of a lifetime. A sampling of episodic memories would include remembering your current psychology professor's name, what you had for dinner last Tuesday, the color of your first “two-wheeler,” and your mother's maiden name. From the standpoint of research on memory, the list of words you just learned and recalled in an experiment is an episodic memory. The critical aspect here is that the memories are part of your personal history and are not generally shared by others.

Semantic memory, conversely, refers to your general world knowledge, including your knowledge of the vocabulary and rules of language, and the general knowledge that relates concepts and ideas to one another. For instance, you know what a bird is, you know how to use the word “bird” in a sentence, and you know that robins and sparrows are typical of the category, whereas ducks and penguins are less so. While it's obviously true that much of your semantic memory knowledge originally came from personal experience, the particulars of that experience tend to be either forgotten or irrelevant. You probably don't remember the first time someone said to you “That's a bird. It's a kind of animal that flies.” And even if you do, that memory is fairly unimportant when you glance out the window, see a bird, and effortlessly realize that it's a robin.

While it's true that episodic and semantic memory are highly interdependent, we will nonetheless devote this chapter almost exclusively to episodic memory principles and processes. These principles and processes are the more traditional issues in the study of human memory, including phenomena such as rehearsal, recall, and forgetting. Chapter 6 is devoted to the semantic memory system, its organization and principles of operation. Since semantic memory includes such a vast array of topics, a good deal of the rest of the book is about semantic memory as well. As you'll begin to see, however, the principles and processes investigated in episodic memory research are highly relevant to our later consideration of semantic issues.

**Preliminary Issues**

Let's start our discussion of episodic memory by considering what people know about their own memory systems. This is the topic of metamemory, knowledge about (= meta) one's own memory, how it works, and how it fails to work. While the research and theoretical treatments of metamemory have come largely from developmental investigators, it is clear that the issue is of genuine interest when studying adult cognition as well.

---

1 Metamemory refers to knowledge about our own memory systems. It is the memory component of the more global metacognition, which includes not only knowledge about one's own memory, but also about the functioning of one's entire cognitive system.
Metamemory

Flavell and Wellman (1977) noted that there are “four broad, partially overlapping categories of memory-related phenomena” (p. 3) of interest in studies of memory development. The first category involves basic mental operations and processes, such as recognition and the ability to retrieve information. The mental processes discussed in chapter 3 certainly fit into this category, and some of the chapter 4 material on working memory does as well. The second through fourth categories, they note, correspond to Brown’s (1975) labels of “knowing,” “knowing how to know,” and “knowing about knowing.” These categories reflect a person’s basic storehouse of information in memory (knowing), a person’s knowledge of how to learn or acquire new information (knowing how to know), and a person’s awareness of how the whole memory system functions (knowing about knowing). It is obvious that category two points to an important developmental issue—we learn more and more information as we grow from childhood to adulthood, and the richer knowledge structure in memory makes it easier to store other information as well. What is not so obvious, perhaps, is that a child’s knowledge of how to learn (category three) and a child’s awareness of his or her own mental functioning and processing (category four) grow simultaneously. For our purposes here, these last two categories generally refer to the study of metamemory.

Look back now to Figure 4–12 in the last chapter, from the Kellas et al. (1975a) study. Third, fifth, and seventh graders were given a learning task, and their performance across the serial positions of the list was examined. Our point in that chapter was that recency reflects the operation of working memory, and that Kellas et al. found no differences in the recency effect for these grade levels. Look now, however, at the early portion of the graph, the primacy effect portion. Performance varied a great deal by grade level in this part of the serial position curve. As Kellas et al. pointed out, this strong developmental trend was explainable in terms of the children’s rehearsal. That is, the younger subjects spent less time rehearsing the items as they studied the list, so they were able to recall fewer of them from long-term memory.

What is interesting about this kind of result, as Flavell and Wellman (1977; also Flavell, 1970) pointed out, is that younger children tend to be relatively unaware of the need to engage in deliberate rehearsal. Their performance usually improves dramatically when they are taught a particular rehearsal strategy (Ashcraft & Kellas, 1974; Moely, Olson, Halwes, & Flavell, 1969), so this is not an issue of children using a rehearsal strategy so poorly that it doesn’t help. Instead, the difficulty is at a different level. Children seem to lack an awareness of how to learn; there is an absence of self-monitoring as they perform a learning task (e.g., Leal, Crays, & Moely, 1985). As Flavell and Wellman summarized it: “The term production deficiency was initially coined to describe a child’s failure to use any particular memorization strategy spontaneously when the situa-
tion called for it, even though he could and would use that strategy effectively if explicitly directed to do so by someone else (Flavell, 1970). The sense of the present discussion is that the young child may have a far more general and pervasive ‘production deficiency’ than that: He may seldom think of deliberately trying to retrieve at all, or of deliberately trying to prepare for future retrieval at all, in response to situations that commonly elicit precisely these sorts of cognitive efforts in more mature individuals. Part of metamemory development, then, may consist of coming to know when and why one should intentionally store and retrieve information” (p. 10).

We have two distinct ideas here, both important to our study of episodic memory. One is the notion of self-monitoring, a self-awareness of how well you are doing in a situation that demands memorization or learning. Useful research is now being done on this self-monitoring or awareness factor, giving greater substance to the notion of metamemory processes (cf. Cavanaugh & Perlmutter, 1982). One such study, by Pressley, Levin, and Ghatala (1984), showed that adults favored an associative method over a simple repetition method as they learned new vocabulary words. They only showed this preference, however, after a practice period with the two methods had revealed how inadequate the repetition method was (see also Lovecace, 1984). In contrast, the 11- and 13-year olds not only required this practice, but also required explicit feedback on their performance before showing metamemory awareness. In other words, they seemed to be performing the task blindly, without evaluating how they were doing.

As a general rule, there is an increase across development in the child’s sensitivity to memory and performance and an increase in the self-monitoring and self-correcting that a child does while learning. This increase in awareness then leads the child to better methods of acquiring and remembering information, as well as greater insight into the workings of the memory system itself. Across development then, children become less deficient at producing appropriate strategies for learning and remembering. Part of this improvement is presumably due to the self-monitoring function, noticing that poor performance improves when some action is taken to help learn and remember.

Let’s consider a pertinent example. You know that as an adult you won’t remember material from your text and lecture notes unless you engage in some active, effortful study in preparation for an exam. You take for granted the awareness that you must take action in order to learn the material. Your awareness then motivates you to develop some sort of study plan, some method of making sure that you acquire and remember the course material. You then monitor your performance as you study (or you should, anyway), and correct or modify your study habits as a function of that feedback.

The second notion in Flavell and Wellman’s remark is that of rehearsal, a deliberate and planned strategy for practicing and learn-
ing material that needs to be remembered later. Adults are fairly proficient at generating strategies for remembering, at developing methods to rehearse material they realize will be hard to learn. Further, they aren’t surprised when unrehearsed material can’t be remembered (you’re not really surprised when you can’t remember someone’s name shortly after being introduced, particularly if you made no special effort to remember the name in the first place). Your metamemory processes include the awareness that things don’t merely “get into” memory. You must “get them in” by performing some intentional activity.

This intentional mental activity and its effects on remembering make up the study of rehearsal. To understand the case, a good deal of research in cognitive psychology has examined rehearsal—what it is, how it works, and how it can be applied. Thus, this chapter describes what is known about rehearsal, the major “control process” that Atkinson and Shiffrin (1968; 1971) had in mind in their model of human memory. We’ll start this treatment with probably the oldest form of rehearsal, and the oldest application of metamemory awareness in history—an analysis of mnemonic devices.

Mnemonic Devices

The term mnemonic (the first m is silent; ne-mahn’-ick) means “to help the memory.” It always refers to an active, strategic kind of learning device or method, a rehearsal strategy if you will. Formal mnemonic devices rely on a preestablished set of memory aids and considerable practice on the to-be-remembered information in connection with the preestablished set. Informal mnemonics, such as those you invent yourself, are generally less elaborate, are more suited to smaller amounts of information that you’re trying to remember, and are more idiosyncratic and personalized. The strengths of such mnemonic techniques are many, and include the following important principles: (a) the material to be remembered is practiced repeatedly; (b) the material is integrated into an existing memory framework; and (c) the device provides an excellent means of retrieving the information. We’ll study two of the traditional mnemonic devices first, then turn to the issue of inventing new mnemonics as the need arises.

The first historical mention of mnemonics is in Cicero’s De oratore, a Latin treatise on rhetoric (the art of public speaking, which in Greek and Roman days meant speaking from memory). In this work, Cicero describes a technique based on visual imagery and memorized locations, ascribed to the Greek poet Simonides (circa 500 B.C.). The mnemonic is now commonly referred to as the method of loci (“loci” is the plural of “locus,” meaning “a place”; pronounced low’-sigh). The source of Simonides’ inspiration in inventing the mnemonic technique, as the story goes, was a personal experience. Simonides was performing a lyric poem at a banquet when he was called out of the hall for a message. While he was outside, the roof of the hall caved in. The disaster was so bad that the bodies of the guests, mangled by the falling roof, were unidentifiable. Simonides, however, was able to identify the dead by remembering the places around the banquet table where they had been sitting. Apparently, the visual image of the banquet scene was impressed strongly enough in Simonides’ memory that he was able to remember the faces as they were arranged around the banquet hall locations.3

As a general summary of the method of loci, Cicero wrote: “...persons desiring to train this faculty [of memory] must select places and form mental images of the things they wish to remember and store those images in the places, so that the order of the places will preserve the order of the things, and the images of the things will denote the things themselves, and we shall employ the places and images respectively as a wax writing-tablet and the letters written on it” (from Yates, 1966, p. 2). In other words, decide upon a set of places or locations, a set that can be recalled easily and in order. To apply this mnemonic yourself, you might select a set of 10 or 12 locations you encounter in a walk across campus, or those you encounter as you arrive home. These will be the preestablished memory aids. Now, form a mental image of the first thing you want to remember, and then mentally place that thing into the first location in your set of “loci,” continuing with the second item in the second location, and so forth. Form a good, distinctive mental image of the item in its place (McDaniel & Einstein, 1986); despite intuitions to the contrary, bizarre-ness per se in the images appears to be less important than distinctiveness (Kroll, Schepeler, & Angrin, 1986). When it’s time to recall the items, all you need to do is mentally stroll through your set of locations, “looking” at the places and “seeing” the items you have placed there. Table 5–1 gives an example of this technique.

Another mnemonic device is worth mentioning here as well, partly because it is so commonly known and easy to use. The technique known

3The more impressive mnemonic feat performed by Simonides, reciting the lyric poem from memory, is usually overlooked in psychological accounts of the mnemonic device. Orators in ancient Greece recited in the oral tradition for the obvious reason—portable, written versions of the poems, i.e., books, were not available. The art of rhetoric, or oration, involved extensive memorization, for example, committing to memory the entire Iliad or Odyssey. To make this task somewhat easier, these heroic poems were filled with mnemonic aids, including easily recalled structure or organization, rhyme and meter, images, repeated phrases (“the rows of the great horses” was one of our favorite collocations), and of course a regularly progressing plot. As I recall from freshman literature, the Iliad’s 24 chapters progress “into” the story up until chapter 12, at which point the order of the chapter topics reverse to progress “out” of the story. In the middle chapters, the intricate designs on Achilles’ shield are described; the successive designs essentially summarize the plot of the 12 ordered chapters. Thus, a poet who was reciting the Iliad would call to mind the visual image of the shield, and remember the next part of the story by looking at the next picture or scene in his visual image. We usually simplify our treatment of the classic mnemonic devices as if they were only useful for remembering unrelated sets of words like grocery lists and so forth. At least as they were applied in ancient Greece, the mnemonic devices were immensely more useful than this.
as the peg-word mnemonic (e.g., Miller, Galanter, & Pribram, 1960), in which a memorized set of words serves as a sequence of mental "pegs" on which the to-be-remembered material can be "hung." The peg words rely on rhymes with the numbers one through ten, such as "One is a bun, two is a shoe," and so forth (see Table 5–2). The material being learned is then "hung" on the pegs, item by item, making sure that the rhyming word and the to-be-remembered word form a mental image. For the list "cup flag horse dollar . . . ," create a visual image of a flattened tin cup, dripping with ketchup, inside your hamburger bun; for flag, conjure up a visual image of your running shoes with little American flags fluttering in the breeze as you run a marathon; and so forth (go ahead and form images for the rest of the list as an exercise to understand the principles of mnemonic devices). Now, at recall all you have to do is first remember what peg word rhymes with one, then retrieve the visual image of bun that you created, looking inside to see. . . a cup. Similarly, what peg word rhymes with "two," and what image do you find along with "shoe"?

Bower’s (1970) classic article noted that cognitive psychology had begun to supplement the anecdotal evidence of the usefulness of these mnemonic devices with experimental evidence. To illustrate the effectiveness of such mnemonic devices, he described a study by Ross and Lawrence (1968), in which subjects used a set of 40 campus locations as their loci, then had to learn several 40-item lists using the method of loci. The items were presented about one every 13 seconds, and were followed by an immediate recall test; subjects also returned the next day for a delayed recall test. Average performance on immediate recall, using the method of loci, was 38 out of 40, in their correct order! One day later, subjects averaged 34 correct, again in order. These levels of accuracy are surely a testimony to the effectiveness of such mnemonic techniques.

Bower concluded by noting the two essential ingredients involved in the successful use of the method of loci: first, that imaginal associations are formed between the memorized loci and the words to be learned; and second, the loci are used as cues for guiding recall. In contrast to the subject who merely attempts to learn 40 unrelated words, the subject using the method of loci has a "known bank of pigeonholes or file cabinets in which he stores the list items. At recall, the person knows where to start his recall and how to proceed from one unit to the next; he has a way to monitor the adequacy of his recall; he knows when he has forgotten an item; and he knows when he has finished his recall" (p. 502). The systematic mnemonic device forces you to learn the material well, provides a memorable, durable record in memory, then guides you during retrieval.

**Mnemonic Principles**

Reread the last sentence in the previous section. It contains the crux of the argument for mnemonic effectiveness. What three things does a mnemonic device do to help the memory? First, it provides a structure for learning, for acquiring the information. The structure may be relatively elaborate, as a set of 40 loci would be, or it may be simple, e.g., rhyming peg words. It may even be highly arbitrary if the material is not particularly extensive. (The mnemonic for the names of the five Great Lakes, remembering the word HOMES for Huron, Ontario, Michigan, Erie, and Superior, isn’t especially related to the to-be-remembered material, but it is quite simple.) Second, by means of visual images, rhymes, or other kinds of associations, the mnemonic ensures a durable record of the material in memory, one that won’t easily be forgotten (what’s sticking out of your running shoes?). As such, the mnemonic seems to safeguard against interference in storage, or against
other kinds of loss within memory itself. And finally, the mnemonic guides
you through retrieval by providing effective cues for recalling the infor-
mation. As we'll discuss extensively later on in the chapter, this func-
tion of the mnemonic device is critically important, since much of what we
call 'forgetting' seems often to be a case of retrieval difficulty.

This three-step sequence should sound familiar to you since it's essen-
tially the same sequence we mentioned in the definition of memory in
chapter 1. Memory can be defined as the initial acquisition of information,
the retention of that information, then finally the retrieval of the infor-
mation (Melton, 1963). Logically, your performance in any situation that
requires memory is dependent on all three of these steps. Any one of the
three might be the faulty process that accounts for poor performance, and
all three must be accomplished successfully for good performance. From
the standpoint of mnemonic devices, it is important that you use system-
atic mnemonic methods to guarantee success at each of the three stages.
If you apply the method of loci correctly, you will acquire the information
at the outset, it will remain in memory, and it will be retrievable. If you
conscientiously follow the mnemonic scheme for remembering people's
names, you will be able to remember those names (see the Suggested
Readings at the end of this chapter for sources that describe these and
other mnemonic devices).

In a moment, we'll turn to the cognitive psychology of this three-step
process, the empirical research on episodic memory. If you understand
this section on mnemonic devices, then you'll be prepared for that
research—you'll already understand, at an intuitive level, many of the
memory principles that mnemonics rely on. Before turning to this discus-
sion, however, let's briefly consider the issue of self-generated mnemonics.

**Invented Mnemonics** As a college student, you are continually
faced with the need to learn and remember new information. Sometimes
this information is quite new to you, and is difficult to learn because it is
so unrelated to things you already know. A first course in chemistry or
computer programming, for example, asks that you learn to think about
new concepts, using an unfamiliar vocabulary. At other times, the new
information you must learn is only an amplification or elaboration of
things you already know; German III, after having taken I and II, or Cog-
nitive Psychology, after taking Introductory Psychology. How do you deal
with these memory demands?

Start first with what you know about metamemory. How aware are you
of your own state of knowledge about new information? Do you spend
more time studying when the material is unfamiliar? Do you monitor your
performance as you read a text, testing your memory as you go along, or
do you merely read through a chapter, treating all the information in the
same, undifferentiated way? Flavell and Wellman's review indicated that
we become increasingly sophisticated in our abilities to monitor our own
memory performance, and in our abilities to invent ways of learning new
material. While this is certainly the case, based on the developmental lit-

erature, it is also obviously true that everyday adult memory could stand
some improvement too. (Did you notice, in the Pressley et al., 1984, voca-

bulary learning study that even the adults needed the practice session
before realizing that simple repetition was inferior to associative elabo-
ration?) Attention to the self-monitoring, self-correcting aspects of meta-
memory is vital here. You should be able to point to aspects of your per-
formance that signal this kind of metamemory awareness—a slower
reading rate on unfamiliar material, more note taking on difficult lecture
topics, and so forth. If you can't find any behavioral way to prove your
metamemory awareness, then you probably are as "production deficient" or
complex memory tasks as small children are on simple ones.

Secondly, when you are consciously aware of the need for special effort
or difficult material, what do you do? Do you "think about knowing," that
is, think of how difficult it will be to learn the material and what you might
do to actually learn it? Do you ever invent your own mnemonic devices to
make sure you'll remember something? Recall the HOMES example and
how a simple acronym or word can function as a retrieval cue for the
needed information. Other principles come into play here as well, espe-
cially rhymes and rhythm; to use a common example, if I ask you how
many days there are in June, don't you rely on the "30 days hath Septem-
ber" mnemonic?

As a lesson to yourself, pick something from this book to learn by an
invented mnemonic, maybe the seven themes listed in chapter 2, maybe
the three characteristics of automatic processing described in chapter 3.
Try to come up with a mnemonic based on images, rhymes, or acronyms
so that you can remember the list easily. The very act of inventing the
mnemonic will probably ensure that the material is stored strongly in
memory, of course, and the image or rhyme will help assure that it
remains in memory until you test yourself. Make sure your mnemonic
itself is memorable, so that it can guide you through retrieval by cueing
the several items you're trying to remember. 4 For a really impressive dem-
stration, start applying the method of loci to a large body of informa-
tion, maybe the overall content of this chapter (if it worked for Simonides,
surely it can work for you). Finally, be on the lookout for "ready-made"
mnemonics, or situations in which a mnemonic association is almost built
into the material. (Did you notice in chapter 3 that Cherry did research
on the cocktail party phenomenon, and can't you "see" that maraschino
cherry in the cocktail glass?) The more you notice such easily associated
items, the easier it will be to come up with invented mnemonics for other
material as well.

---

4 Ms. Monica Castelli, a student in my Memory and Cognition course, devised an excellent men-
monic for remembering the seven themes I listed in chapter 2. The initial letter of each of the words
in the mnemonic sentence stands for one of the themes. Her sentence was: "Always amazing—study-
ing data related to memory," for the themes Attention, Automatic/conscious processes, Serial/par-
allel processing, Data versus conceptually driven processes, Representation of knowledge, Tacit
knowledge and inference, and Metacognition.
The Ebbinghaus Tradition of Memory Research

Let's turn now to the mainstream research on human learning and memory, on how people acquire and remember new information. We'll begin this story at its proper beginning, with the unparalleled model of research given us by the first serious human memory investigator, the German psychologist Hermann Ebbinghaus.

As indicated in chapter 1, the tradition of scientific research on human memory was begun over 100 years ago, when Hermann Ebbinghaus published his monograph Über das Gedächtnis (1885; Memory: A Contribution to Experimental Psychology is the English translation, first published in 1913, then reprinted in 1964). Ebbinghaus, as you'll recall, disagreed with the conventional wisdom of the time—that the higher mental processes could not be studied scientifically, or that they could only be studied through the method of introspection. He consequently set about to devise a method by which reliable, scientifically acceptable evidence about human memory could be obtained. As is commonly known, Ebbinghaus used only himself as a subject in his studies, but was careful to note the circumstances of his testing and other procedural details so that his own results have largely been replicated by later researchers, using more “acceptable” procedures. In the process of his investigations, he had to invent his own memory task, his own experimental stimuli, and his own set of procedures for testing and data analysis. Few, if any, could do as well today in the absence of guidance from research and colleagues. As Hilgard put it, in his introduction to the 1964 edition, “For the beginner in a new field to have done all of these things—and more—is so surprising as to baffle our understanding of how it could have happened” (p. vii).

Furthermore, Ebbinghaus anticipated a variety of important issues in human memory research, not to mention developments in general scientific methodology. For instance, Nelson (1985) notes that the relearning task Ebbinghaus invented was a “radical idea that was far ahead of its time, both methodologically and conceptually” (p. 472), in that it recognized the possible influence of nonconscious factors (recall how committed those of Wundt's school were to the dissection of the structure of the conscious mind). Ebbinghaus carefully laid out his method of measuring retention of information, and thus anticipated Bridge's (1927) influential notion of the operational definition (Nelson, 1985, p. 474). In devising ways to analyze his numerical results, he even came close to inventing what we now would call a correlational- or within-groups t test (Ebbinghaus, 1885/1964, footnote 1, p. 67).

Among the psychological issues that Ebbinghaus investigated were the time course of forgetting, the length-difficulty relationship, and an explanation of serial (ordered) learning. Probably more important than these contributions, as Murdock (1985) observes, were the general methodological developments Ebbinghaus was responsible for. We tend to think of Ebbinghaus, simplistically, merely as the inventor of the nonsense syllable, the “meaningless” consonant-vowel-consonant (CVC) triads like XIG or BEF, which he used as stimuli in all of his studies. As you'll see, this is a seriously impoverished view of Ebbinghaus's pioneering contributions.

To begin with, it is instructive to consider why Ebbinghaus felt compelled to invent a “meaningless stimulus” to be used in his studies of memory. His rationale was that he wanted to study the properties of memory and forgetting—the fundamentals. It was clear to him that words, if used as experimental stimuli, would hopelessly complicate his results. Had he acquired and remembered a list of items by the simple exercise of memory, or had his performance been altered by his existing knowledge of the words? Putting it simply, learning seems to imply acquiring new information using whatever mental processes are required. Yet words are not new, so “learning” a list of words in some sense involves a misnomer.  

Other features of Ebbinghaus's procedures are important to note as well, since they exhibit the care and forethought he devoted to his work. For instance, his insight led him to control for possible mental and physical changes, by testing himself at the same time of day; for different degrees of learning, by setting a fixed learning criterion of one perfect recitation without any hesitations (the criterion was two such recitations in some studies); and for the possible intrusion of mnemonic or nonrote factors, by adopting a rapid presentation rate of 2.5 items per second. His only experimental task was the relearning task, in which a list is originally learned, set aside for some period of time, then later relearned to the same criterion of accuracy. The savings score derived from this task was the reduction, if any, in the number of trials (or the time, in other studies) necessary for relearning, compared to original learning. Thus, if a list required 10 original learning trials, but only 6 upon relearning, there was a 40% savings (4 fewer trials on relearning divided by the 10 original trials).

As noted previously, the invention of this task revealed an amazing sensitivity to fundamental states of memory. That is, Ebbinghaus wanted to study memory per se, not just the conscious memory of past events that is measured by a recall task, but even seemingly unrecallable memories that might influence performance. By the method of relearning, any informing...
of the day in this study, in contrast to the situations in which the time of
day was kept constant).

Some other fundamental results obtained by Ebbinghaus, impressive
not necessarily because they are surprising, but because they were the first
empirical demonstrations of the effects, are nearly as familiar. For ex-
ample, he investigated the effects of repetitions, studying one list 32 times
and another 64 times, where the 32-trial condition approximated the stan-
dard learning criterion of one perfect recitation. Upon relearning, the
more frequently repeated list showed about twice the savings of the less
frequently repeated list—in other words, overlearning yields a stronger
record in memory. Longer lists were found to require more trials to learn
than shorter lists, certainly an unsurprising result; the more interesting
result was that longer lists then showed higher savings upon relearning.
In essence, while it is harder to learn a long list originally, the longer list
is then remembered better, simply because there was more opportunity
for overlearning it (there were more trials in learning before eventual mas-
tery of the whole list). In fact, in one experiment (Chap. 8), Ebbinghaus
continued to relearn the same set of lists across a five-day period. The
savings scores he obtained showed a trend that, if extrapolated, would
eventually show perfect savings, i.e., no forgetting at all. As an interest-
ing contrast here, Ebbinghaus also reported his results on relearning pas-
sages of poetry (kept at 80 syllables in length). After the fourth day of learning,
the savings was 100%. (Is it any wonder that actors overlearn their parts
through multiple rehearsals?)

**Evaluating the Ebbinghaus Tradition**

There is no doubt or disagreement whatsoever that Ebbinghaus had a
tremendous impact on what came to be known as the field of Verbal
Learning, and by extension, cognitive psychology too. In a recent set of
papers commemorating the 100th anniversary of the 1885 publication (see
Suggested Readings), a consensus began to emerge concerning the value
of Ebbinghaus’s work (with some passionate dissenters, of course). In
Mandler’s (1985) words, Ebbinghaus’s contribution represented “a door
being opened into the human mind, the realization—contrary to then
established wisdom—that it is in fact possible to gain positive knowledge
about human memory” (p. 464). Slamecka (1985a) called him “the
founder of our discipline,” and summarized his influence by saying: “He
set out to show that an empirical science of memory was possible. . . . He
succeeded admirably in this enterprise . . .” (Slamecka, 1985b, p. 497).

Disagreement does exist, however, on the balance between Ebbing-
haus’s positive and negative influences. No one claims that the general
methodological aspects of his work exerted a negative influence; the care
he took in controlling extraneous factors, the meticulous way he investi-
gated the effects of variables, and the example he set of inventing an

---

**FIGURE 5-1**

The classic forgetting curve from Ebbinghaus (1885). The figure shows the reduction in
savings across increasing retention intervals, time between original learning and
relearning.
objective task with which to assess memory were all tremendously positive influences.

On the other side of the coin, however, he did exclude meaning from his studies, after all, by developing nonsense syllables and then testing his memory for them almost exclusively. Much subsequent research, up through the '60s and even the '70s, continued to use nonsense syllables as the items to be learned and remembered (see chapter 1, where Glaze's 1928 tabulation of the 'meaningfulness' of nonsense syllables is discussed). The problem with this, it is generally conceded, is that for the most part people will not deal with a truly meaningless stimulus. Instead, they will attempt all sorts of mediating, mnemonic, or other rehearsal strategies to render a 'nonsense' syllable sensible—mentally turning BEF into BEEF for example. Ebbinghaus seemed able to avoid such mnemonic strategies partly because of the rapid rate of presentation and partly because of his desire to study only rote learning. It is doubtful that most subjects in psychological research can behave in that pure a fashion, however. Indeed, the fact that Noble's (1952) subjects found their rating task sensible (judging "How meaningful is this nonsense syllable?") confirms that people will attribute or even invent meaning if none is there.

It is clear from Ebbinghaus's own discussion that the nonsense syllable was a deliberate effort on his part to avoid the complications of uncontrolled meaningfulness. In Slamecka's (1965a) analysis, this was understandable and even praiseworthy when we consider that Ebbinghaus was the first to attempt to deal with any of these issues in a scientific way. Contemporary criticism of the artificiality of laboratory research (see the Neisser quotation at the beginning of this chapter) asserts that Ebbinghaus's example misled psychology into excluding meaningfulness from its research, and inspired decades of irrelevant, ungeneralizable, and inapplicable results. To quote Kintch's (1985) evocative remark, "What a terrible struggle our field has had just to overcome the nonsense syllable! Decades to discover the 'meaningfulness' of nonsense syllables, and decades more to finally turn away from the seductions of this chimera. Instead of the simplification that Ebbinghaus had hoped for, the nonsense syllable, for generations of researchers, merely screened the central problems of memory from inspection with the methods that Ebbinghaus had bequeathed us" (p. 461).

The Ebbinghaus tradition, in short, has been understood (or misunderstood) as an admonition, to the effect that 'meaning complicates matters, so eliminate it from the stimuli.' A more temperate view might note that, in the absence of previous research, Ebbinghaus quite properly simplified the experimental situation so as to get interpretable results, a view that Kintch (1985) also stated. We might further suggest that the failure lies less with Ebbinghaus than with his successors, who slavishly stuck to his methods without questioning their intent or usefulness.

The Current Position. If we consider the quite different position that is accepted today, and contrast it with the model of Ebbinghaus's research, we arrive squarely at the empirical research on episodic long-term memory. Today's position consists of at least three parts. First, the fact that people invent meaning, regardless of the experimenter's wishes, is taken as evidence that human memory relies very heavily on meaning. Our memory is 'set up,' loosely speaking, to learn in terms of the information's meaning to us. To put it bluntly, if people deal so heavily in meaning, then perhaps we should investigate how they invent meaning rather than try to prevent them from doing it.

The second element of this position is implied by the first, that the subject in our memory experiments is an active participant, not content to recite syllables passively and have them eventually make an impression on memory, as Ebbinghaus did, but instead is intent on applying mental resources and strategies to virtually every learning situation. The active subject assumption, in other words, is quite the opposite from the Ebbinghaus model, although he himself was aware of the tendency (after all, he took great pains to prevent such activity).

The third part of the current position, implied by some of the critics mentioned before, is that results based on meaningless stimuli are themselves meaningless when we attempt to understand how people learn and remember. This point, in essence, claims that results based on memory for meaningless information can't possibly tell us how normal memory works, because there is such a strong influence of meaning on how people remember information. This is the issue of ecological validity again, saying in essence that our traditional laboratory results do not apply to real-world situations that involve memory for meaningful material. As an example, we might analyze what someone remembers after reading a passage from a book. The elements of meaning in that passage will exert a strong influence on what is remembered, as will the person's own knowledge of the material. Our understanding of that act of remembering, it is said, will be terribly inaccurate if limited to Ebbinghaus-inspired principles based on nonsense syllable experiments.

Another aspect of Ebbinghaus's methods that became 'standard operating procedure' has largely escaped criticism, which is strange in view of the current popularity of the ecological validity issue. Ebbinghaus paced himself at 2.5 items per second during learning; the overwhelming majority of subsequent studies, even those conducted now, have maintained this 'experimenter-paced' aspect of the learning task. A strong conviction of my graduate school mentor was that we learn more about human learning if we let subjects pace themselves through a list they are trying to learn; see the description of the Kellas et al. study that accompanies Figure 4-12, for instance. Kellas's procedures were to record unobtrusively the amount of time a subject spent studying each item in the list, then to observe the dependency of recall on study time. While there are some interpretive complexities involved in this method (as there are when meaningful stimuli are used, or when learning is experimenter-paced), it is certainly a sensitive method for understanding subjective-controlled factors such as rehearsal. And it clearly bears much more relationship to real-world learning episodes—after all, how often do you try to learn a list of things by spending only one or two seconds per item? (See, for instance, a paper by Johnson & Kieras, 1983, where self-paced study of technical material brought these subjects up to the level of another group that was already familiar with the material.)

I find no inherent difference in the status of the nonsense syllable issue and the experimenter-paced learning issue, at least in terms of generalization to more realistic learning situations. How peculiar that our memory studies still routinely use this part of Ebbinghaus's technique, while critics rail against the 'meaningless stimuli' aspect of his procedures.
Storage of Information in Episodic Memory

How do people store information in episodic memory, the long-term memory system for personally experienced events and information? How is new information recorded in this long-term memory system so that it will be preserved until some future time when it is needed? And how can we measure this storage of information? Ebbinghaus’s research investigated one principle kind of storage variable, repetition, and one memory task, relearning. He found that an increase in the number of repetitions led to a stronger memory, a trace of the information in memory that could be relearned more quickly. This would suggest that frequency is a fundamental variable in learning—information that is presented more frequently will be stored more strongly in memory. (As a quick example, do you remember from chapter 1 the year that the first psychological laboratory was established by Wundt? The date was intentionally repeated several times, to permit this kind of example.) A corollary of this idea is that people should be good at remembering how frequently something has occurred. Indeed, Hasher and Zacks (1984) have summarized a large body of research that shows how sensitive people are to the frequency of events, so sensitive in fact that these authors propose that frequency information is automatically encoded into memory, without deliberate effort or intent (but see Greene, 1986).

We will consider three important “storage” effects here, rehearsal, organization, and imagery. A summary of these three will then lead us to the topic of retrieval and a discussion of forgetting.

Rehearsal

A fundamental statement on storage was made by Atkinson and Shiffrin (1968) in their influential model of human memory. In their formulation, information that resides in short-term memory may be subjected to rehearsal, a deliberate recycling or practicing of the contents of the short-term store. Atkinson and Shiffrin proposed that there are two effects of rehearsal. First, rehearsal maintains information in the short-term store, preventing it from being lost or displaced by other information. Second, the longer an item is held in short-term memory by rehearsal, the greater the probability that the rehearsal will also store the item in long-term memory. Basically, this position states that rehearsal “copies” the item into long-term memory, with the strength of the long-term memory trace depending on the amount of rehearsal. In short, rehearsal transfers information into long-term memory (see also Waugh & Norman, 1965). Of course, in most experimental situations, the items being transferred are words that the subject already knows. Thus, “transferring information” is generally taken to mean storing some “tag” or other indication that a certain word was an item in the list being learned.

What evidence is there of this transfer function for rehearsal? Aside from the classic Ebbinghaus study on repetition, many experiments have shown that rehearsal of information leads to better long-term retention. For example, Hellyer (1962) used the Brown-Peterson task to examine the effects of rehearsal. Subjects were shown a CVC trigram, as usual, and were asked to perform an arithmetic task between study and recall, also as usual. The difference in this study was that on some trials the trigram had to be spoken out loud once, on some trials twice, four times, or eight times. Figure 5-2 shows the results of this experiment. The more frequently rehearsed the item was, the better it was retained across the distracting period of arithmetic. Using a somewhat different approach, Hebb (1961; cited in Loftus & Loftus, 1976) presented strings of digits for recall. Whereas performance remained stable and low on those strings that changed from trial to trial, performance improved significantly on the string of digits that was periodically repeated in the learning trials. Although the task seemed to be testing only short-term retention of digit strings, mere repetition of a string led to some transfer of the information into long-term memory.

While these studies and many others confirmed the general notion that rehearsal leads to an improvement in long-term memory performance, the evidence they presented was somewhat indirect. After all, for evidence about long-term memory, it was slightly odd that two short-term memory

\[ \text{FIGURE 5-2} \]

Hellyer’s (1962) results, recall accuracy as a function of the number of rehearsals afforded the three-letter nonsense syllable and the retention interval.
paradigms had been used, the Brown-Peterson task and a digit recall task. More direct evidence on rehearsal was soon to follow, however. Among the many such studies, the standard citation on rehearsal is to a series of studies performed by Rundus (1971; Rundus & Atkinson, 1970).

Rehearsal and Serial Position Effects In his experiments, Rundus had his subjects learn 20-item lists of unrelated words, presenting them at a rate of 5 sec per word. Subjects were asked to rehearse out loud as they studied the lists, repeating whatever words from the list they cared to during each 5-sec presentation. Rundus then tabulated the number of times each of the words had been rehearsed, and compared this tally to the likelihood of recalling the word correctly in the free recall task. Figure 5–3 shows his most telling results. In the early primacy portion of the serial position effect there was a direct positive relationship between the frequency of rehearsal and the probability of recall. In fact, Rundus also examined a proportional measure of rehearsal and found that “For a given amount of rehearsal, items from the initial serial positions have no better recall than items from the middle of the list” (Rundus, 1971, p. 66). In other words, the primacy effect—higher recall of the early items—was viewed as entirely dependent on rehearsal; the early items can be rehearsed more frequently (no doubt because of the experimenter-paced task; see footnote 8), so are then recalled better.

Think for a moment about how this research fits in with the serial position effects we discussed in the last chapter. In that section, several variables were shown to have an effect only on the recency portion of the serial position curve, the portion generally viewed as dependent on short-term memory. For instance, requiring subjects to count backward by threes for a period of time seemed to destroy their short-term memory record of the last items in the list. Importantly, this manipulation did not alter the primary portion of their recall, the portion due to long-term memory. Conversely, a different set of variables left the recency effect alone, but had a big effect on primacy. For instance, increasing the study time per item from one to two seconds improved recall of the early positions, but did not change recall from the end of the list. Thus, the typical U-shaped serial position curve (showing that “memory sags in the middle,” as Martin and Noreen quipped in 1974) is generally taken to indicate two kinds of memory performance; primacy effects due to recall from long-term memory, and recency effects due to short-term memory recall. As a logical extension, Rundus’s results showed that deliberate, overt rehearsal improved the accuracy of recall in primacy, the portion of the list influenced by long-term memory. No such effect was found on the recency items, however.

Later research has focused on the rehearsal activity per se, rather than on the general issue of whether rehearsal stores information in long-term memory or not. Kellas, McCauley, and McFarland (1975b), for instance, compared a group of subjects who rehearsed out loud, as Rundus had tested, to one that rehearsed silently. They found that the group that rehearsed out loud had a tendency to rehearse and recall the items in order, as if they had been required to perform a serial recall task. The silently-rehearsing subjects, however, took advantage of the free recall instructions and recalled the words in a more typical order, last items first and then the initial items. Kellas et al. noted a further difficulty. The subjects in the vocal rehearsal condition took longer during study to achieve the same level of accuracy as the silent rehearsal subjects. This suggested that rehearsing out loud deprived those subjects of their flexibility in rehearsal; having to rehearse out loud led them to the rather unimaginative strategy of “repeating the words out loud.” In contrast, subjects in the silent rehearsal group were probably able to use more elaborate and complex kinds of rehearsal, for instance, devising sentences or images to help remember the words. These elaborate kinds of rehearsal couldn’t be easily vocalized, Kellas et al. reasoned, so they disappeared in the groups that rehearsed vocally. It is apparently not as easy to externalize the mental process of rehearsal as was originally believed.

Two Kinds of Rehearsal

The theoretical position behind the notion of “more elaborate and complex” rehearsal claims that there are two major kinds of rehearsal, each with different effects on storage (Craik & Lockhart, 1972). According to
this position, maintenance rehearsal is a low-level, repetitive kind of information recycling. This is the kind of rehearsal you’d use to recycle a phone number to yourself until you dial it. The essential idea here is that, once you’ve stopped rehearsing the information, it has left no permanent record in memory at all. In Craik and Lockhart’s view, maintenance rehearsal (or Type I rehearsal) merely maintains information at a particular level in the memory system, without storing it more permanently or deeply. As long as an item of information is subjected to maintenance rehearsal, it can be retrieved. Once the maintenance rehearsal stops, however, the item should vanish without a trace.

Elaborative rehearsal (or Type II rehearsal), on the other hand, is a more complex kind of rehearsal that uses the meaning of the information to help store and remember it. When information is subjected to elaborative rehearsal, according to Craik and Lockhart, the information is stored more deeply in the memory system, at a level that makes contact with the meaning of the information. As a consequence, material that was rehearsed elaboratively should be more permanently available for retrieval from memory—in short, should be remembered better. Among other things, you might include imagery or mnemonic elaboration in your elaborative rehearsal; you might try to construct sentences from the words in a list you’re trying to learn; you might impose some organization or structure on the list; you might even try to convert “nonsense syllables” like BEF into more meaningful items, like BEEF. Thus, maintenance rehearsal maintains an item at its current level of storage, whereas elaborative rehearsal moves the item more deeply into the memory system.

Craik and Lockhart (1972) advanced a notion of the memory system quite different from the “stage” approach of sensory, short-, and long-term memory that we’ve become accustomed to. They embedded their proposal of two kinds of rehearsal into what they termed a levels of processing, or depth of processing framework. Since this framework prompted a flurry of research activity, we need to elaborate its description a bit, and understand some of the research that has supported it. After all, the possibility that rehearsal, even if only of the “maintenance” variety, might not improve long-term memory performance at all is rather surprising.

Depth of Processing

The essence of the depth of processing framework goes as follows. Any perceived stimulus receives some amount of mental processing. Some stimuli, which receive only incidental attention, are only processed to a very “shallow” level in memory, possibly no deeper than a rather sensory level (as in hearing the sound of the words without attending to meaning, as a daydreamer might do during a lecture). Other stimuli, on the other hand, are subjected to more intentional and meaningful processing. This deeper processing elaborates the representation of that item in memory, for example, by drawing relationships between already-known information and the item currently being processed. Thus, superficial kinds of processing, which require little attention to the meaning of an item, correspond to maintenance rehearsal, or “shallow encoding” into memory; Nairne (1983) has drawn a strong parallel, in fact, between Type I rehearsal and the articulatory loop of working memory. Meaningful processing, which requires much more attention and effort, corresponds to elaborative rehearsal, a “deep encoding” of the material. An important theme in this depth of processing framework, notice, is that the mental activities a subject engages in during processing are as important for understanding memory as a determination of the “final resting place” of the information (what we’ve termed short- or long-term memory).

Several predictions from the depth of processing framework have been tested with a fair degree of initial success. For example, if information is processed shallowly, with only maintenance rehearsal, then the information should not be particularly memorable on a later recall test—if it was only maintained, then it should not have been stored in long-term memory by that maintenance (technically, according to the Craik and Lockhart position, we should say that it is not stored at a deep, meaningful level). An interesting confirmation of this prediction was reported by Craik and Watkins (1973), in a task that was expected to lead to only shallow, superficial processing. The subjects in this experiment heard a long series of words on a tape recording. Their task was to monitor the words, listening for those that began with some critical letter, say G. When a new “G-word” was presented, they were to remember the new one, and forget the previous one. At the end of the list, subjects merely had to report the last G-word they had heard. This procedure was followed across several trials, with each trial having a different critical letter. Because of the way the lists had been constructed, Craik and Watkins could specify exactly how long each of the critical words had been maintained by Type I rehearsal. Table 5–3 gives an example of a list like those used by Craik and Watkins, and illustrates what they termed the “I-value”, the number of intervening items across which a critical word was maintained.

As you might expect, at the end of the set of trials, subjects were given a surprise recall test, in which they were to recall any and all words they had heard from all the lists. As you might not expect, the amount of time an item had resided in short-term memory (the I-value) had no effect on the subjects’ recall. Craik and Watkins did find that memory was better for lists that had been presented at a slower rate, in agreement with earlier research, but this effect was independent of simple time in short-term memory. As they put it in their conclusion, “This result is clearly contrary to the idea that recall probability necessarily increases in direct proportion to the total amount of time an item has been thought about... time in short-term store does not by itself lead to long-term retention...
Challenges to “Depth of Processing”

Now that you understand the Craik and Lockhart notion of depth of processing, it’s time to pull the rug out from under you—enthusiasm for the depth of processing approach has dimmed considerably in recent years. Much of this dimming has been due to Baddeley’s (1978) important review paper “The Trouble with Levels.” We’ll focus on two important aspects of this review, the problem of defining “levels” independently of retention scores, and the influence of task effects (see papers by Glenberg, Smith, and Green, 1977, and Glenberg and Adams, 1978, for a discussion of a related problem, defining and manipulating Types I and II rehearsal).

Defining “Levels”  To begin with, there is a lot that appeals to common sense in the levels of processing framework. After all, we are all intuitively aware of situations where we heard or read some material (lecture, textbook), but then later seemed not to have retained any of the information at all. In retrospect, we realize that our attention was not engaged while listening or studying, and that our processing of the information was quite superficial. In particular, we may have noted some of the sounds of the lecturer’s words, that is, the phonological aspects, or possibly would have noticed a change in the written or orthographic form of the text (boldface, italics). Nonetheless, we weren’t processing the meaning of the stimuli, but just the shallow, physical features (Glenberg and Adams, 1978, in fact, found that Type I rehearsal was particularly likely to involve phonetic-acoustic information). Naturally, a recall test on the meaning will show poor retention.

Common sense, however, is a poor basis for building a science of memory. As Baddeley (1978) noted, a critical difficulty with the levels of processing approach was its rather circular definition of “depth.” That is, material that was remembered well was said to have been processed “deeply,” down to the level of meaningfulness and semantic elaboration. But how can we tell if some task requires deep or shallow processing? The answer is “Test it, and if recall is high, the task must have required deep processing.” Obviously, the difficulty with this definition is that there is no way of deciding independently how deeply a stimulus is processed or encoded, except by looking at performance. To illustrate this circularity, consider a hypothetical experiment on elaborative rehearsal. To demonstrate that elaborative rehearsal leads to better performance, for instance, we ask people to rehearse in a certain meaningful way, then test their recall. First, we need to assure ourselves that they used elaborative rehearsal. Did the subjects rehearse deeply? Yes, because their recall was high. Well, then, does deep, elaborative rehearsal improve long-term memory? Yes, because their recall was high.

Task Effects  Since the time of Ebbinghaus, we have known that different memory tasks will shed different kinds of light on the variables that affect performance. Ebbinghaus used a relearning task instead of simple recall, so that even material that was difficult to retrieve might have a chance of influencing performance. In a similar vein, a substantial difference is generally found between performance on recall versus recognition tasks. In recognition tasks, subjects are shown items that were originally studied, known as “old” or “target” items, as well as items that were not on the studied list, known as “new” or “distractor” items. They must then decide which items are targets and which are distractions. Accuracy on a recognition task is usually much higher than on a recall task. (As a simple example of this difference, consider which of the following is easier, recalling a specific term as a “fill-in-the-blank” answer, or picking the term as one of four multiple choices? The fill-in item is a recall task and multiple choice is a recognition task.)

The reason that recognition is easier, it is generally agreed, is that recognition tasks require much less retrieval effort than recall tasks—indeed, recognition doesn’t seem to require deliberate retrieval at all, since the to-be-retrieved answer is presented to the subject, who then only has to make a “new/old” decision. Since more information is stored in memory than can be easily retrieved, recognition generally shows greater sensitivity to the influence of stored information. The relevance of this effect to the issue of depth of processing, simply, is that most of the early research that
supported the levels of processing approach relied on recall tasks. A distinct possibility, then, was that recognition tests might prove more sensitive in testing the effects of Type I and Type II rehearsal.\footnote{In fact, even standard recognition memory tasks may be somewhat insensitive. Jacoby and Dallas (1981) found that regardless of whether subjects could recognize a passage as one they had read before, those passages were read more rapidly the second time. This effect suggests that there may be a lingering memory trace for information even if subjects have no conscious recollection of having experienced the information before. See Kalter and Roediger (1984) and Graf and Schacter (1987) for a discussion of this “memory without awareness” phenomenon, also termed explicit vs. implicit memory. Moscovitch, Winocur, and McLachlan (1986) present an interesting application of this principle to testing memory deficits in elderly adults with Alzheimer’s disease.}

The clever set of studies by Glenberg et al. (1977) can be taken as a definitive confirmation of this possibility, and as one of the more serious challenges to the depth of processing position. Glenberg et al. used a standard Brown-Peterson task, but asked subjects to remember a four-digit number as the (supposedly) primary task. During the variable length retention intervals, subjects had to repeat either one or three words out loud as a distractor task (don’t confuse the distractor task here with “distractor items,” items tested in recognition that were not shown originally). Since the subjects were led to believe that digit recall was the important task, they presumably devoted only minimal effort to the word repetitions; i.e., they probably used only maintenance rehearsal or Type I processing.

At the end of the 60 experimental trials of Experiment 1, the subjects were surprised with a free recall test on the words they had spoken during the distractor periods. Consistent with the levels of processing view, recall of the words showed no effect of the varying amount of rehearsal, that is, of the period of time (2, 6, or 18 seconds) during which subjects had repeated the words out loud. In Experiment II, however, the final memory test was a recognition task. Words that had been repeated during the retention intervals were mixed together with words not used in the experiment, and subjects had to indicate if they had or had not encountered each of the words in the experiment. In this surprise test, the rehearsal interval did influence performance—words rehearsed for 18 seconds were recognized significantly better than those rehearsed for shorter intervals. To ice the cake, Experiment III showed the same effects of rehearsal interval in the recognition task, but this time using CVC triads instead of meaningful words; again, the longer an item had been rehearsed, the better it was recognized. (Given what you know about nonsense syllables and about maintenance rehearsal, why was this study “icing on the cake”?)

Other researchers have shown the influence of task in a slightly different way. A study by Morris, Bransford, and Franks (1977) will illustrate this effect. Subjects in levels of processing experiments are typically given an orienting or acquisition task to perform on the set of stimuli; for example, one group might be asked to mark each “a” in the list of words, another to check off the words rhyming with “boat,” and another to check off the members of the “animal” category. These tasks, presumably, require processing that ranges from rather shallow or superficial levels to deep, meaningful levels. Morris et al. continued this tradition by having some subjects decide whether two words rhymed, and other subjects decide if a word fit a supplied semantic “frame” (see Table 5-4 for examples of the experimental conditions; the term “orienting task” simply implies that the task oriented subjects toward one or another kind of processing, in this case semantic vs. rhyme processing). When tested with a recognition task that was sensitive to semantic factors, subjects who had performed the semantic-orienting task had higher scores (84%) than those who had acquired their information in the rhyming task (63%). Importantly, of those subjects tested with a recognition task sensitive to rhyming factors, those who originally performed the rhyme-orienting task did better than those in the semantic-orienting group (49% vs. 33%). Of course, in the depth of processing view, semantic processing should always lead to better retention, since semantic coding processes information more elaborately in memory. The Morris et al. study provides a pointed counterexample—only when semantic retention is measured will semantic coding be superior. Under other circumstances, e.g., a different kind of recognition test, more “shallow” processing might result in superior performance.

**On Balance** How should we view the depth of processing idea? Baddeley (1978) concluded that, all things considered, the depth notion is valuable only at a rough, intuitive level, and that other, more detailed approaches deserve careful study rather than the broad and general principles that Craik and Lockhart attempted to identify. We agree with Baddeley (1978) that a scientific approach to memory must do better than

---

**Table 5-4**

**EXPERIMENTAL CONDITIONS IN MUKHS, BRANSFORD, AND FRANKS (1977)**

<table>
<thead>
<tr>
<th>Semantic Orienting Task</th>
<th>Word</th>
<th>Decision</th>
<th>Recognition Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Does the word make sense?</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The _______ had a silver {$\text{engin}$}.</td>
<td>Train</td>
<td>Yes</td>
<td>0.644</td>
</tr>
<tr>
<td>The _______ had a silver {$\text{engin}$}.</td>
<td>Eagle</td>
<td>No</td>
<td>0.859</td>
</tr>
<tr>
<td>Rhyme orienting task</td>
<td><em>Does the word rhyme?</em></td>
<td>Eagle</td>
<td>Yes</td>
</tr>
<tr>
<td>_______ rhymes with legal</td>
<td>Peach</td>
<td>No</td>
<td>0.524</td>
</tr>
</tbody>
</table>
provide intuitive definitions and circular logic. And we agree with the conclusions of the experiments that semantic coding will be superior only when a semantic test is given, and that the superiority of elaborative rehearsal may only be supported in recall, as opposed to recognition tests.

On a more everyday level, however, we might do well to remember Baddeley’s description of the approach as a rule of thumb, and look for potentially valuable implications. That is, we might think of maintenance and elaborative rehearsal, loosely, as corresponding to simple recycling in short-term memory, on the one hand, and more complex, meaningful study and transfer into long-term memory, on the other. As a rule of thumb for evaluating your own learning, this certainly gives you a decent idea of what will and what won’t improve your memory. When you read a text, do you merely process the words at a fairly simple level of understanding, or do you elaborate what you’re reading, searching for connections and relationships that will make the material more memorable (as in the “icing on the cake” question above)? Add to this the following observation. While recognition performance does show the effects of even maintenance rehearsal and activity, are you going to be given a recognition test on your exam? If your exams are recall tests, then you should probably bear in mind the advantages of elaborative rehearsal.

**Organization in Storage**

Another vitally important piece of the “storage puzzle” involves the role of **organization**, the **structuring or restructuring of information as it is being stored in memory**. Part of the importance of organization is derived from the powerful influence it exerts—well-organized material can be stored and retrieved with impressive levels of accuracy. Another part of its importance (at least in my view) is that it helps maintain the topic’s organization of information, to create the critical link between memory and understanding. One way of looking at this is that we are not passive recipients of stimulation, but rather, are active participants in learning situations, intentionally seeking ways of making information more memorable.

The earliest program of research on organization (or clustering) was conducted by Bousfield. In his earliest study (Bousfield & Sedgewick, 1944), Bousfield had asked subjects to name, for example, as many birds as they could. The intriguing result was that the subjects tended to name the words in subgroups, for instance, “robin, bluejay, sparrow—chicken, duck, goose—eagle, hawk.” To investigate this further, Bousfield (1953) designed a study in which subjects were given a free recall task, with a 60-item list to be learned. Unlike other work at that time, however, Bousfield used related words for his lists. In particular, the 60-item list was made up of 15 words each from four distinct categories: animals, personal names, vegetables, and professions. The subjects were presented these list items in a randomized order, yet in free recall, they tended to write the items down by category; for instance, “dog, cat, cow, pea, bean, John, Bob,” etc.

Now, how can we explain this, sensible though it may be? How is it that subjects recalled the words in an order that was so different from the input order, and furthermore, in an order that obviously reflected the relationships among the different words? Bousfield’s most immediate interpretation was that the greater-than-chance grouping of items into clusters “implies the operation of an organizing tendency” (p. 237).

Where did this “organizing tendency” come from? It was not in the words themselves, to be sure. While the words were drawn from categories that were assumed to exist in the subjects’ memories, it would be foolish to say that the words exerted the tendency to organize themselves. But the tendency was in the **subjects**, in their unspoken mental activities that went on during the learning of the list. Obviously, subjects noticed at some point during input that several words were drawn from the same categories (given an order like “dog, bean, John, cat, pea, Bob,” an attentive subject might notice the structure in the list upon the first repetition of a category, at “cat”). As a reasonable strategy for doing well on such a long list, they elected to group the items together on the basis of category membership (notice the metamemory effect implied here as well). This implies that subjects were reorganizing the list by means of rehearsal as it was presented; for instance, a subject seeing “cat” at the fourth position would then rehearse it in conjunction with “dog” because of their common category. The consequence of this reorganization during storage would be straightforward—the way the material had been stored governed the way it was recalled (another possibility, that reorganization occurs during recall instead of during storage, will be discussed later).

Investigations of category clustering became very common following Bousfield’s initial reports for several important reasons. A widely shared viewpoint, expressed neatly by Mandler (1967) was that: “memory and organization are not only correlated, but organization is a necessary condition for memory” (p. 328, emphasis added). On this view, any information that was stored in memory was, almost by definition, organized. Further, standard storage strategies, most prominently rehearsal, came to be viewed as organizational devices with the consequence that anything rehearsed was also organized (at least, anything rehearsed elaboratively). Mnemonic devices, in this view, were no different; for instance, “all organizations are mnemonic devices” (Mandler, 1967, p. 329), and likewise, all mnemonic devices provide organization. We will discuss two of the reasons for the rash of studies on clustering here, along with examples of the research that supported those reasons. Chapter 6, which covers semantic memory, delves still more deeply into the topic of organization. Indeed, the clustering research, with its focus on how word meaning affects recall, was probably the most important experimental bridge to studies of long-term semantic memory.
Recall Within each recalled category, items occur in the order of recall. Memory for each category is tested in a separate block of trials. The number of items in each category is the same across blocks and trials. This design allows for an examination of the effect of category on memory performance. The recall order is randomized across blocks and trials to control for any effects of order. Figure 4-3 shows the results of a repeated measures ANOVA on the number of items recalled in each category. The main effect of category was significant, indicating that participants recalled more items from certain categories than from others. The interaction between category and block was also significant, suggesting that the difference in recall performance across categories varied across blocks.

Figure 4-3

Table 4-2: Average Number of Words Recalled Over Four Trials

<table>
<thead>
<tr>
<th>Category</th>
<th>Trials 1</th>
<th>Trials 2</th>
<th>Trials 3</th>
<th>Trials 4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>7.0</td>
<td>6.9</td>
<td>7.1</td>
<td>7.2</td>
<td>28.4</td>
</tr>
<tr>
<td>B</td>
<td>7.3</td>
<td>7.2</td>
<td>7.1</td>
<td>7.4</td>
<td>29.4</td>
</tr>
<tr>
<td>C</td>
<td>7.6</td>
<td>7.5</td>
<td>7.7</td>
<td>7.6</td>
<td>30.4</td>
</tr>
</tbody>
</table>

**Figure 4-4**

The figure shows the relationship between the number of words recalled and the number of trials. The data points are plotted on a scatter plot, with each point representing a participant. The line of best fit is also shown, indicating a positive correlation between the two variables. The correlation coefficient is 0.75, suggesting a strong relationship. The regression equation is y = 0.5x + 2, where y is the number of words recalled and x is the number of trials.
tified by rehearsal frequency. This result, of course, extends the repetition-recall relationship originally observed by Ebbinghaus.

Confirmation of the notion of hierarchies came from both the recall order results and the retrieval time measure. In particular, recall of the items within a category was quite rapid, with only about .5 seconds between words. Going from the end of one category to the beginning of the next was a considerably longer process, taking about 1.5 seconds at the beginning of the list (see also Patterson, Meltzer, & Mandler, 1971). The important idea here, according to most investigators, was that recall was a process of accessing the hierarchy, then decoding the higher-level units into the individual items (refer to Figure 5-4, in which a decoding of “minerals,” then “metals,” then “common” would lead to the items “aluminum,” “copper,” etc.). Finding evidence of this hierarchical organization in the retrieval order and time scores was particularly compelling. Thus, this study illustrated, as did a number of others, the close relationship between rehearsal and organization, and the powerful effects of these on recall.

**Subjective Organization** To close this section on organization, note that we have been discussing clustering only in the context of lists that contain members of common conceptual categories—animals, vehicles, and so forth. It would be a serious mistake to imply that organization is relevant only in such circumstances, however. Organization is presumed to be imposed by the subject, by the active mental reorganization that occurs during rehearsal, regardless of the inherent structure of the material. This reorganization is most easily studied with structured, categorized lists, but is not limited to such lists. Tulving (1962) provided a clear demonstration of this principle, which he termed **subjective organization**. Subjective (or subject-defined) organization represents the same kind of organizing and structuring of a list of items, but without the experimenter-supplied category structure.

Tulving used the multitrial free recall task, presenting the same list of words across several trials, each trial showing a new reordering of the words. His analysis looked at the regularities that developed in the subjects' recall orders, i.e., the consistent groupings of otherwise unrelated words that occurred trial after trial in the task. As an example, a subject might recall the words “dog, apple, lawyer, brush” together on several successive recall trials. This consistency, despite the experimenter's reordering of the list items from trial to trial, suggested that the subject had formed a cluster or chunk composed of those four items based on some idiosyncratic basis. For example, a subject might link the words together in a kind of sentence or story: “The dog brought an apple to the lawyer, who brushed his hair.” Regardless of how these links were formed, the clusters were then used repeatedly during the recall trials, serving subjectively as the same kind of organized unit that “dog, cat, horse, cow” would in an experimenter-defined cluster. Thus, Tulving pointed out that even “unrelated” lists become organized because of the mental activity of the subject who imposes this organization. As he put it: “Perhaps paradoxical this suggests that a list of completely unrelated words is probably as fictional as is a truly nonsensical nonsense syllable” (p. 352).

**Imagery**

The last storage variable to be considered here involves **visual imagery**, the mental picturing of a stimulus that then affects later recall or recognition. Of course, we’ve discussed two prominent visual imagery effects already, the mental rotation studies, which suggest an imaginal code in working memory, and the imagery-based mnemonic devices. What we are focusing on now, however, is the effect that visual imagery has on the storage of information into long-term memory, the possible boost that imagery gives to material you’re trying to learn.

The name most closely associated with research on imagery effects is Alan Paivio. In his book, Paivio (1971) reviewed scores of studies that illustrated the generally beneficial effects of imagery on learning and retention. These beneficial effects are over and above those due to other variables, such as word- or sentence-based rehearsal, or meaningfulness (e.g., Bower, 1970; Yuille & Paivio, 1967).

As just one example, Paivio described a **paired-associate learning** study by Schnorr and Atkinson (1969). In such a study, items are presented in pairs, the first item designated the “stimulus” item, the second the “response” item. The subject's task here is to learn the list so that the correct response item can be reproduced whenever the stimulus item is presented. Thus, if you saw the pair “elephant-book” during study, you would be tested during recall by seeing the term “elephant” and your correct response would be “book” (the later section on interference describes this task in more detail). Schnorr and Atkinson had their subjects study half of a paired-associate list by means of imagery, forming some visual image of the stimulus and response terms together. The other half of the list was studied by means of rote repetition. On immediate recall, the pairs learned by imagery were recalled at better than 80% accuracy, compared to about 40% for the rote repetition pairs. The superiority of the imagery condition was found even after a one-week retention interval.

Studies such as this one led Paivio to propose the **dual coding hypothesis** (e.g., Paivio, 1971). This hypothesis states that words that denote concrete objects, as opposed to abstract words, can be encoded **into memory twice**, once in terms of their verbal attributes and once in terms of their imaginal attributes. Thus, a word like “book” enjoys an advantage in memory studies. Because it can be recorded twice in memory: once as a word and once as a visual image, there are two different ways it can be retrieved from memory, one way for each code. A stimulus term like “idea,” on the other hand, probably has only a verbal code avail-
able for it, since it does not have an obvious imaginal representation. (This is not to say that people can’t eventually create an image to help remember a word like “idea,” but merely to say that the image is much more available and natural for concrete words.)

There is much more to be said about imagery, and in general, other nonverbal means of representing and remembering information (Kosslyn, 1978, 1981; Paivio, 1971). Some of this research is discussed in the next chapter, where we consider semantic memory and the role of imagery in general world knowledge. For now, consider an episodic memory study by Watkins, Peyserigoi, and Brems (1984). These investigators tested the idea that rehearsal, most commonly studied with verbal materials, might also have a pictorial or imaginal component. They presented picture-word pairs to two groups of subjects; one group was given verbal rehearsal instructions, the other group was given pictorial rehearsal instructions (“try to maintain an image of the picture in your mind’s eye”). The pictures were displayed by slide projector, the words by tape recorder. Some of the picture-word items were followed by a 15-second interval of time, during which rehearsal could take place, and some were immediately followed by the next picture-word pair, thus preventing rehearsal. Figure 5-5 shows three sample pictures from the study (top row).

After the list of pairs had been presented, subjects were given a cued recall task, in which either a fragment of the printed word or a fragment of the picture was presented as a cue; examples are shown in the middle and bottom rows of Figure 5-5. Accuracy in identifying the fragments depended on all three factors: type of rehearsal, type of fragment, and rehearsal interval. As Figure 5-6 shows, the pictorial rehearsal group outperformed the verbal rehearsal group when the picture fragments were presented during the test, but only if the picture had been followed by the 15-second rehearsal interval (top half). Not surprisingly, when picture rehearsal subjects were given the word fragments during the test (bottom half), their performance did not improve when an opportunity had been provided to rehearse. Both types of rehearsal, in other words, improved performance when the recall test was consistent with the subjects’ type of rehearsal. Neither type improved performance if the retrieval cues were inconsistent with the type of rehearsal.

**Storage Summary · Encoding Specificity**

Storage is just one side of the coin, so to speak. The other side, retrieval, is obviously just as important—what good does it do to store something in memory if you can’t retrieve it when you need it? Let’s conclude this section on storage with a short summary that also previews several important ideas for the topic of retrieval. What common threads are there in the studies of episodic memory storage? What generalizations can we draw from research on rehearsal, organization, and imagery? How are we to understand the phenomenon of storage into episodic memory? The best way to understand storage, it would seem, is to consider it in light of retrieval, in terms of how the material will be retrieved from the memory system.

Think back to the research you’ve just read about. What variables or conditions led to the best performance? There were certainly some experiments in which subjects were able to recall information *despite* their activity during learning—Hebb’s subjects recalled the repeated digit list, for instance, and Glenberg et al.’s subjects recognized words they had originally been told were only part of a distractor task. Indeed, Ebbinghaus’s results, with the incredibly rapid presentation rate of 2.5 items per second, attest to the ability of the human memory system to acquire information under adverse conditions (see Tulving, 1985), a kind of “brute force” memory effect. But, for the most part, the research shows that memory performance is enhanced when there is a close correspondence between the study and test conditions, in other words, when the activities during testing match those of acquisition.
As a simple example of this principle, try naming the 12 months of the year in alphabetical order. This “thought experiment” doesn’t need to be confirmed by laboratory data. It’s obvious that the way you have acquired and stored the information, the way it’s organized, is based on chronological order. Retrieving the information in such an *incongruent* fashion, one that mismatches your storage organization so thoroughly, is quite difficult. A related example makes the underlying principle even clearer. I occasionally run into a former student “out in the world,” that is, out of the school context in which I met the person. When this happens, I usually can’t remember the person’s name, even though I can often remember the name when the student comes to my office. The mismatch between storage and retrieval contexts turns me into an “absentminded professor,” at least in that circumstance.

Now consider again the Morris et al. (1977) study. Subjects who made semantic judgments during acquisition did well on a semantic recognition task; subjects who had made rhyme judgments did well on a rhyme recognition task. Under the other conditions, though, the mismatch or *incongruity* between acquisition and retention tasks yielded poorer performance. Likewise, Watkins et al. (1984) found that incongruous conditions, picture rehearsal with word-fragment cues, and vice versa, led to poorer performance than congruent conditions, i.e., the picture-picture and word-word conditions.

Something that occurs during acquisition facilitates memory performance when retention is tested in a way that matches acquisition. Congruence, or agreement between acquisition and retention conditions, improves our performance. Incongruence between acquisition and retention depresses our performance. Why?

In Tulving and Thompson’s (1973) view, an important reason is captured by the phrase *encoding specificity*. By this phrase, Tulving and Thompson meant that information, the list you’re trying to learn for example, is encoded into memory not as a set of isolated, individual items. Instead, *each item is encoded into a richer memory representation, one that includes any extra information about the item that was present during encoding*. Thus, when you encounter “cat” in a list of words, you are likely to encode not only the word “cat,” but also related information about that word. Importantly, this extra information would probably include “animal,” in other words, easily retrievable concepts that are semantically related to “cat.” In this situation, “animal” serves as your higher-order label or code for the clustered items “cat,” “dog,” etc.

When your memory is tested, in a free recall task for instance, you attempt to retrieve from memory the record or trace left by your original encoding. In this circumstance, if you encoded “animal” along with “cat,” then “animal” should be an excellent retrieval cue for recalling “cat”—a *useful prompt or reminder for the information to be retrieved*. If you study pictures under a picture rehearsal condition, as Morris et al.’s subjects did, then picture cues will enhance your performance. If the prevail-
ing conditions during the test are quite different, e.g., if you are given word-fragment cues after pictorial rehearsal, then the likelihood of your retrieving the information decreases. The original context cues will provide you with the best access to the information during a recall attempt. If you can’t retrieve those cues, however, your recall suffers.

In summary, storage of information into episodic long-term memory is affected by rehearsal, by organization, and by imagery. The presence of all three of these leads to a stronger memory trace. Further, congruence between study and test conditions seems critical. Relevant rehearsal, including organizational and imaginal elements, improves performance, as does the provision of retrieval cues that were part of the original encoding of the material. Rehearsal that turns out to be irrelevant for the test conditions, however, is generally of little benefit.

Retrieval of Episodic Information

We turn now to the other side of the coin, the retrieval of information from episodic memory. You’ll recall that retrieval and forgetting have been discussed in previous chapters, when we focused on sensory and short-term working memory. You’ll also recall from those sections that two theories of forgetting have preoccupied cognitive psychology from the very beginning, decay and interference theories. Difficulties in the retrieval of sensory memory information were attributed to both decay and interference (e.g., Averbuck & Coriell, 1961; Sperling, 1960). Likewise, Peterson and Peterson (1959), Waugh and Norman (1965), and Wickelgren (1965), among many others, analyzed short-term memory forgetting in terms of simple decay versus more complex interference. These approaches were no doubt inspired by the study of forgetting from long-term memory, a preoccupation that dates back to Aristotle in philosophy, and to the early Verbal Learners in this century.

Decay

It's a bit unusual for the name of a theory to imply the content of the theory so clearly as the term decay. Nonetheless, that's what decay theory is all about—the older a memory trace is, the more likely it has been forgotten, that it has decayed away just as the print on an old newspaper fades into illegibility. Thorndike (1914) enunciated this principle in his law of disuse: Habits, and by extension, memories that are used repeatedly are strengthened, and habits not used are weakened through disuse. Thorndike’s proposal was a beautiful example of a theoretical hypothesis—easily understood, straightforward in its predictions. Unfortunately, it’s wrong, at least as far as long-term memory is concerned.

Interference

A definitive attack on decay theory was provided by McGeoch (1932). In his influential paper, he argued from both theoretical and empirical grounds that decay theory was fundamentally wrong, that time per se is an inadequate basis for understanding or predicting loss of information. Instead, McGeoch claimed that the activities that occur during a period of time are responsible for forgetting, that these activities produce interference that disrupts retention. As Hall (1971) described it, McGeoch’s central point was that “time should be thought of as a conceptual framework within which events take place. The iron bar that is tossed in the field does not rust because of time; rather, oxygen from the air combines with the iron to produce rust, and this process takes place within a time interval.... In short, environmental events themselves, rather than time intervals alone, must be considered in order to account for forgetting” (pp. 459–460).

Interference theory, and tests of interference effects, became a staple in the experimental diet of verbal learning psychologists, much to the current dismay of some cognitive psychologists. At least two reasons for this trend can be stated. First, the argument against decay theory and for an interference approach was convincing, both on theoretical and empirical grounds. Demonstrations like the often-cited Jenkins and Dallenbach (1928) study made complete sense within an interference framework—after identical time delays, subjects who had slept after learning recalled more than those who had remained awake (Figure 5–7). The everyday activities encountered by the awake subjects seemed to interfere with their memory for the list; fewer interfering activities intervened for the sleeping subjects, with the result of better performance.

A second reason for the popularity of interference studies, it seems, is that interference effects were easily obtained in the laboratory, especially with a task already in wide use, the paired-associated learning task. This task seemed “natural” for studying the components of interference. For one thing, it was almost infinitely adaptable—almost any variable thought to influence interference could be incorporated easily. For another, it conformed in an obvious way to the theoretical Zeitgeist or “spirit of the times”; the stimulus-response pairs of items, with a stimulus term coming to “elicit” the proper response, bore a direct and obvious relationship to the stimuli and responses of the animal learning laboratory, and to the dominant theoretical commitment to S-R behaviorism.

Paired-Associate Learning Let’s spend a few moments studying the paired-associate learning task (often simply abbreviated as P-A learning), so you’ll understand interference theory, and so you’ll know one of our tried-and-true methods of testing memory. The basic elements of a P-A learning task are these. A list of stimulus terms is paired, item
to the stimulus terms would grow across repeated trials, showing that (in the terminology of behaviorism) the correct responses had become conditioned to the appropriate stimuli. A typical procedure was to bring a subject up to some predetermined accuracy criterion, say one perfect trial, then ask the subject to learn another list of paired associates.

Table 5–6 presents several P-A learning lists to use as a demonstration. Let’s deal only with list 1 for a moment. Cover this first list with another piece of paper, then slowly move the paper downward, one line at a time. Do this repeatedly, trying to learn the right responses to the stimuli. Keep track of how many trials it takes you to recall all 10 of the response terms correctly.

You’ve now participated in the first half of a typical P-A learning study on interference. The second half of the study involves learning another list of paired associates; the nature or composition of the second list, and its relationship to list 1, is of critical importance. Pick any one of lists 2, 3, or 4 from Table 5–6 and learn it to the same criterion, making sure to note how many trials it takes you to recall all response terms correctly.

If you picked list 2, you should have experienced little or no interference, since list 2 contains stimulus and response terms that are not similar to list 1’s terms. In the terminology of interference theory, you were in the

---

**Table 5–6**  
**LISTS OF PAIRED-ASSOCIATES**

<table>
<thead>
<tr>
<th>List 1</th>
<th>List 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>tall-bone</td>
<td>safe-fable</td>
</tr>
<tr>
<td>pan-leaf</td>
<td>bench-idea</td>
</tr>
<tr>
<td>nose-fight</td>
<td>pencil-owe</td>
</tr>
<tr>
<td>park-flea</td>
<td>wait-blouse</td>
</tr>
<tr>
<td>grew-cook</td>
<td>student-duck</td>
</tr>
<tr>
<td>rabbit-few</td>
<td>window-cat</td>
</tr>
<tr>
<td>pear-rain</td>
<td>house-news</td>
</tr>
<tr>
<td>mass-crowd</td>
<td>card-nest</td>
</tr>
<tr>
<td>print-kiss</td>
<td>color-just</td>
</tr>
<tr>
<td>smoke-hand</td>
<td>flower-jump</td>
</tr>
</tbody>
</table>

---

A-B—C-D condition, where the letters A through D refer to different lists of stimulus or response terms. The A-B—C-D condition represents a kind of baseline condition in terms of interference, since there is no similarity between the A-B and the C-D terms (note, however, that you may have needed fewer trials on the second list because of “general transfer” effects from list 1, warm up or learning to learn). If you picked list 3, there should have been “massive” negative transfer; in other words, it would have taken you many more trials to reach criterion on the second learning list. The reason for this is that the same stimulus and response terms were used again, but in new pairings. Thus, your experience on list 1, forming strong enough associations to reach criterion, interfered with the later activity of learning to re-pair the items. The standard term for this type of list was A-B—A-Br, where the r stood for “randomized” or “re-paired” items. Finally, if you picked list 4 (the A-B—A-B condition), there should have been considerable positive transfer, requiring many fewer trials to reach criterion on the second list. This is because the response terms in list 4 (designated B', read “B prime”) are highly related to the earlier ones you learned (B).

It should be clear that the transfer effects we’ve been discussing are “proactive,” that is, they demonstrate effects on a current learning task due to a prior one. Of course, we discussed proactive interference, PI, and release from PI, at some length in the last chapter. Table 5–7 gives the general experimental design for a proactive interference study as well as for a retroactive interference (RI) study. To repeat, retroactive interference exists when some learning experience interferes with recall of an earlier experience—the newer memory interferes backward in time (“retro”).

Both proactive and retroactive interference were examined extensively (or excessively, depending on your viewpoint), with complex theories built on P-A learning results. Although a massive literature is available, no attempt will be made to cover it in depth here (but see standard works such as Postman & Underwood, 1973; Underwood, 1957; Underwood & Schulz, 1960; and Klatsky’s very readable summary, 1980, chapter 11). Instead, let’s examine one particular mechanism that was postulated by interference theorists—unlearning—and the difficulties this mechanism encountered when researchers broadened their horizons to include preexisting associations.8

**Problems of Meaning** Briefly, a subject who first learns an A-B list and then must learn an A-C list will experience proactive interference while learning the A-C list (and will experience retroactive interference if list A-B is tested again). This interference was thought to be due to a process of unlearning (e.g., Underwood & Postman, 1960)—learning a list requires the breaking or unlearning of the previous associations (Kintsch, 1970). As investigators began to use real words instead of CVC nonsense syllables in their experiments, the unlearning principle was extended to preexperimental associations, that is, to associations that already existed in the subjects’ memories before the learning experiment was conducted.

A study by Slamecka (1966) illustrated the difficulties in this “jump” into the world of meaningful words. Subjects were first given a list of words, call it A, and were asked to respond with the first association that occurred to them; for example, given “cat,” you might have responded “dog,” given “tree,” you might have responded “bush.” These free associates were recorded so they could be tested later. The A list of words was then used in an A-B list learning task, which was then followed by a second list learning task, using list A-C. Table 5–8 diagrams the three conditions of the study and the three kinds of lists that were tested (the P stands for the list of preexisting associations, like “dog,” which people gave at the beginning). According to the unlearning hypothesis, the original free associates (P) given to the A terms (“cat-dog”) should have been unlearned when subjects learned A-B (“cat-pin”). Likewise, the newly learned B terms should be unlearned during acquisition of list A-C (“cat-desk”).

### Table 5–8

<table>
<thead>
<tr>
<th>Group</th>
<th>Preexperimental Associations</th>
<th>Learn List 1</th>
<th>Learn List 2</th>
<th>Recall</th>
<th>Recall</th>
</tr>
</thead>
<tbody>
<tr>
<td>II</td>
<td>A-P</td>
<td>A-B</td>
<td>A-C</td>
<td>A-P</td>
<td>A-B, A-C</td>
</tr>
<tr>
<td>III</td>
<td>A-P</td>
<td></td>
<td></td>
<td>A-P</td>
<td>A-B</td>
</tr>
<tr>
<td></td>
<td>cat-dog</td>
<td>cat-pin</td>
<td>A-C</td>
<td>cat-desk</td>
<td></td>
</tr>
<tr>
<td></td>
<td>tree-bush</td>
<td>tree-salt</td>
<td></td>
<td>tree-dish</td>
<td></td>
</tr>
<tr>
<td></td>
<td>car-truck</td>
<td>car-book</td>
<td></td>
<td>car-paper</td>
<td></td>
</tr>
<tr>
<td></td>
<td>etc.</td>
<td>etc.</td>
<td></td>
<td>etc.</td>
<td></td>
</tr>
</tbody>
</table>

8The term “preexisting associations” was a theoretical construct; nowadays, these associations would simply be called “semantic knowledge.”
A final recall test was administered to all groups. Subjects in Group I showed 33% lower recall of the B terms than subjects in Group II, because the second group had not participated in the A-C learning condition. Thus, there was evidence for unlearning the B terms because of interference from A-C. On the other hand, there were no group differences at all in subjects’ ability to recall their own original free associates, A-P. What should have been altered by interference, the preexisting associations, was totally immune to it. In the face of existing associations in memory, laboratory-induced interference seemed downright puny.

What were these preexisting associations? They were language-based associations, connections between words (like “cat-dog”) that were firmly entrenched in memory because of years of experience. Why wouldn’t these associations behave, why wouldn’t they demonstrate the kind of interference that, by rights, they should have demonstrated? Maybe they wouldn’t follow the theory of P-A learning because the theory was wrong, or at least inappropriate. That is, memory performance was assumed to follow the laws of learning, and these laws were being investigated with the P-A learning task. Finding so major a difference between laboratory and real world “items,” between CVCs and words, suggested an unsettling conclusion— theories based on traditional learning tasks and paradigms were irrelevant, or at least inadequate, for understanding memory of words and language. It seemed entirely possible that the P-A learning laws weren’t general, that they only applied fully to nonsense syllable learning in a P-A task. As Jenkins (1974) put it in the title of his article: “Remember that old theory of memory? Well, forget it!”

**Retrieval Failure**

Since the mid-1960s, a very different theory has come to dominate cognitive psychology’s view of forgetting. Both the decay and interference theories suggested that information in long-term memory can truly be forgotten, that is, lost from memory. This more technical definition of the term forgetting, loss from memory, was implicit in the mechanisms thought to account for forgetting. For example, “unlearning” was merely a synonym for “extinction,” and implied the actual weakening and loss of learned associations (interestingly, the principle of spontaneous recovery was also tried as an explanation for some interference effects). The current view makes a radically different claim, that in essence, there is no genuine forgetting from long-term memory (save for possible loss due to organic or physical factors, e.g., stroke). Instead, so-called forgetting is viewed as a failure of retrieval, the final member of Melton’s (1963) triology of acquisition, retention, and retrieval.

An early indication of this current view was provided in a study by Tulving and Pearlstone (1966). In this study, two groups of subjects studied the same list of 48 items, four each from 12 different categories (animals, fruits, sports, etc.; other subjects learned shorter lists, or lists with fewer items per category, but we’ll focus only on the two most dramatic groups here). The items were preceded by the appropriate name of the category, e.g., “crimes-treason, theft; professions-engineer, lawyer,” but subjects were told that they only had to remember the items themselves. Because both groups were treated identically until the beginning of the recall period, it can be assumed that both had acquired the same amount of information from the list, and both had retained equal amounts. At recall, one group was asked for standard free recall. The other group was asked for free recall, but was provided the names of the categories as retrieval cues.

The results were predictable, but profound in their implications. The free recall group was able to recall 40% of the list items, while the cued recall group named 62% of the items. In short, the free recall group had only recalled a portion of the items that were learned. We know they learned more than they recalled, because the cued group had been exposed to the same learning and retention conditions, yet had recalled more. As Tulving and Pearlstone put it, “information about many words must be available in the storage … even when this information is not accessible” (p. 389; emphasis added) under free recall conditions.

One conclusion we can draw from these results, which confirms intuitions dating back to Ebbinghaus, is that the recall task often underestimates the amount of information that was learned. Savings scores, not mention scores on recognition memory tests, usually show much higher retention than recall scores. A much more important implication, however, is that unsuccessful retrieval, say in the absence of cues, might prove to be a critical component of forgetting. In fact, retrieval failure might be the major or even the only cause of forgetting. On this view, information stored in long-term memory remains there permanently, so is available just as a book on the library shelf is available. Successful performance, however, also depends on accessibility, the degree to which information can be retrieved from memory. Items that are not accessible are not immediately retrievable, just as the misshelved book in the library cannot be located or retrieved. This position suggests that information is not lost from memory, but instead is lost in memory, so to speak. This loss of access will persist until some effective retrieval cue is presented, some cue that “locates” the item that can’t be retrieved.

**Retrieval Cues and Encoding Specificity**

You already know how access can be increased to an inaccessible memory trace, although you learned that principle under a different name, encoding specificity. The way to increase your access to information in memory is to reinstate the original learning context, to maximize the con-
gruence between current test conditions and the conditions that prevailed during acquisition. In short, access is increased by effective retrieval cues. Any cue that was encoded along with the learned information should increase the accessibility of that information. This is why the category cues helped Tulving and Pearlstone's subjects to recall more than they otherwise would have been able to. Similarly, this is why recognition memory tests usually reveal higher performance than recall tests. In a recognition test, you merely have to pick out which of several alternatives is the correct choice. What better retrieval cue for some piece of information could there be than the very information you're attempting to retrieve?

Subsequent research has demonstrated the power of the encoding specificity principle in quite dramatic fashion. (By far the most convincing demonstration I've ever seen is presented in Tables 5–9 and 5–10, taken from Bransford & Stein's 1984 book on problem solving; do that demonstration now, before reading further.)

Thomson and Tulving (1970) asked their subjects to learn a list of words for later recall. Some of the list words were accompanied by "cue words" printed in lowercase letters; subjects were told they need not recall the cue words, but that the cues might be helpful in learning the items. Some of the cue words were high associates of the list items, for instance hot-COLD, and some were low associates, for instance, wind-COLD. During recall, subjects were tested for their memory of the list under one of three conditions, low- or high-associate cues, or no cues at all. The results were exactly as predicted from the encoding specificity principle. High associates used as retrieval cues benefitted the subjects' recall both when the high associate had been presented during study and when no cue word had been presented. Presumably, when no cue word had been presented, subjects spontaneously retrieved the high associate during input, and encoded it along with the list item. In contrast, when low associates had been presented during learning, only low associates functioned as effective retrieval cues. High associates used as retrieval cues were no better for these subjects than no cues at all. In other words, if you had studied wind-COLD, receiving "hot" as a cue word for COLD was of no value. Encoding specificity thus can even override existing associations during a recall attempt. (Note that encoded cues do not cause "unlearning" of the preexisting association; they simply function as more effective cues during the task.)

More surprising than this, encoding specificity can even override the usual advantage that recognition tests show over recall tests. A series of influential papers by Tulving has demonstrated a paradoxical result, termed "recognition failure of recallable words" (e.g., Tulving & Thomson, 1973; Watkins & Tulving, 1975). In these studies, a weakly associated cue is presented along with the target word during original learning—say "glue-CHAIR." When a recognition test is presented later, the target

This demonstration experiment illustrates the importance of retrieval cues. You'll need a blank sheet of paper and a pencil. Please follow the instructions exactly.

Instructions: Spend 3 to 5 seconds reading each of the sentences below, and read through the list only once. As soon as you are finished, cover the list and write down as many of the sentences as you can remember (you need not write "can be used" each time). Please begin now.

- A brick can be used as a doorstop.
- A ladder can be used as a bookshelf.
- A wine bottle can be used as a candleholder.
- A pan can be used as a drum.
- A record can be used to serve potato chips.
- A guitar can be used as a canoe paddle.
- A leaf can be used as a bookmark.
- An orange can be used to play catch.
- A newspaper can be used to swat flies.
- A TV antenna can be used as a clothes rack.
- A sheet can be used as a sail.
- A boat can be used as a shelter.
- A bathtub can be used as a punch bowl.
- A flashlight can be used to hold water.
- A rock can be used as a paperweight.
- A knife can be used to stir paint.
- A pen can be used as an arrow.
- A barrel can be used as a chair.
- A rug can be used as a bedspread.
- A telephone can be used as an alarm clock.
- A scissors can be used to cut grass.
- A board can be used as a ruler.
- A balloon can be used as a pillow.
- A shoe can be used to pound nails.
- A dime can be used as a screwdriver.
- A lampshade can be used as a hat.

Now that you've recalled as many sentences as you can, turn to Table 5-10.

CHAIR is often not recognized if it appears in a very different context, for instance, in the set, "desk, top, chair." In other words, subjects fail to identify CHAIR as a word they have seen previously in the experiment, since its current context is so different from the original encoding. Following this recognition failure, subjects are then given a cued recall test. Here they routinely do recall CHAIR when presented with "glue" as a retrieval cue. While these experiments used rather arbitrary contexts (who would spontaneously think of "glue" as a cue for CHAIR?) the similarity of the result to the earlier point about congruous study and retrieval contexts is obvious. In short, even simple recognition depends on encoding specificity.
Autoassociative Memories

From Tolman's original experiments, it was clear that rats could learn to navigate complex environments by forming mental maps of the environment. He referred to these mental maps as "cognitive maps," which are not literally representations of the physical environment but instead are mental constructs that animals use to guide their behavior. Tolman believed that these cognitive maps are stored in long-term memory and can be retrieved and used to solve new problems, even in the absence of the physical environment.

In the context of Tolman's work, Ebbinghaus's finding that the amount of new information as a function of time from the original learning session was important. Ebbinghaus's results showed that forgetting is a gradual process, with the rate of forgetting slowing down over time. This finding was significant because it suggested that memories are not erased completely but rather become less accessible over time.

Tolman's research on cognitive maps provided a new framework for understanding memory and its role in navigation. His work challenged the prevailing view that memory was a passive process and suggested that it is an active, constructive process that involves forming mental representations of the environment.

Table 5-10

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Memory Type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Implicit</td>
<td>Knowledge acquired and stored without conscious effort.</td>
</tr>
<tr>
<td>2</td>
<td>Explicit</td>
<td>Knowledge that can be consciously remembered and retrieved.</td>
</tr>
</tbody>
</table>

Tolman and his colleagues (1921) demonstrated the encoding specificity principle:
FIGURE 5-8

Results obtained by Bahrick et al. (1975) in their study of memory for faces and names across 50 years. The two curves that decline dramatically across intervals are recall curves, plotted against the right axis. The curves that depict recognition performance, plotted against the left axis, show remarkably high performance across a long span of time, and begins to decline after about 35 years.

to decline noticeably until about 15 years later, and picture recognition remained in the 80% to 90% range until about 35 years later. And, as Bahrick et al. pointed out, the decline in the very oldest group may have been influenced by factors related to physical aging, possibly introducing a negative bias for the oldest group.

What leads to such impressive levels of retention, particularly when we compare them to the relatively lower performance of subjects in laboratory memory studies? As Bahrick et al. noted, in the typical situation, individuals have learned the names and faces of their classmates across a four-year (or longer) period. This situation is termed prolonged acquisition. According to the authors, this principle has two important components associated with it, overlearning and distributed practice. Both are of critical importance to their results, and both yield useful guidelines for improving your memory.

First, the information tested in the Bahrick et al. study was overlearned, in fact to a much higher degree than laboratory studies have examined (even Ebbinghaus didn't test the effects of a four-year long learning phase). The result of such overlearning is obvious, much-improved retention. Of course, others have found such high retention levels on laboratory tasks. For example, Shepard (1967) found 98.5% accuracy in a picture recognition task, impressive because the subjects were tested on 600 pictures (and Standing, 1973, found 73% recognition on 10,000 pictures). Nonetheless, we wouldn’t imagine that such a high rate would be maintained for very long, based on forgetting results beginning with Ebbinghaus. Indeed, Shepard’s subjects’ recognition accuracy declined to 92% after three days, 87% after one week, and 58% after four months. Thus, for Bahrick et al. to find such accurate recognition performance up to 35 years after original learning was quite amazing, to say the least. It certainly advises us to consider overlearning when life-long retention is the goal. (Imagine being able to recognize the names of psychological principles, hypotheses, and investigators at 90% accuracy on your exams right now, much less 35 years from now!)

Secondly, prolonged acquisition represents learning that was distributed across a very long period of time. In contrast, typical memory experiments have a rather short acquisition period, in which the learning takes place in a fairly “massed” way, i.e., not distributed across intervals of
time, but massed together over a short period. Standard laboratory research has claimed for a long time that distributed practice leads to much better retention than massed practice (e.g., Underwood, Keppel, & Schulz, 1962). Indeed, one of the soundest bits of advice that cognitive psychology gives to students is to distribute your practice and learning, rather than massing (better known as “cramming”) it together; for example, study one hour a night for five nights, rather than five hours on one night. The Bahrick et al. results suggest that the laboratory-based effect is not only general to more naturalistic settings, but it is greatly magnified when naturalistic, everyday memories are tested. In confirmation, Smith and Rothkopf (1984) found that correct recall in a statistics class was 13% higher for distributed versus massed presentation of videotaped lessons.

An increasing number of research studies are now appearing on the topic of genuine autobiographical memory, long-term memory for “real-world” as opposed to laboratory materials. Continuing his earlier work, Bahrick (1984) has reported a study subtitled “Fifty Years of Memory for Spanish Learned in School.” This study, as the previous one, found that retention was much better for this information than for the usual laboratory-learned materials. And, like the previous study, the retention effects were attributed to the level of original learning—the more strongly the information was originally acquired, the better the very long-term retention (Bahrick uses the term *permanence* to refer to this very long-term memory; see also Bahrick, 1983, for a report on memory of a city’s streets and locations 50 years later).

Other autobiographical memory studies are finding a variety of interesting effects as well. For example, Winograd and Killinger (1983) have examined the “flashbulb memories” (Brown & Kulik, 1977) of college students for a significant event, the assassination of President Kennedy in 1963 (note that subjects were being asked to recall their own particular circumstances when news of the event reached them, not whether they remembered the event itself). While their data show an increase in the amount of recallable information as a function of the subject’s age in 1963, they also found evidence that the surprise or shock involved in such events may not be necessary for high levels of retention—subjects showed high recall for the Nixon resignation and the moon landing of the U.S. astronauts, neither of which was an unexpected, surprise occurrence. Distinctiveness of the event, however, is surely important (e.g., Schmidt, 1985). This is more than a little reminiscent of the old-fashioned von Restorff effect, improved retention for a list item that is made distinct or different from the rest of the list, say, by underlining it in red (e.g., Cooper & Pantle, 1967).

And finally, two modern-day Ebbinghauses have adopted the procedure of testing their own memories in carefully controlled, long-term studies; the difference from Ebbinghaus’s procedure was that Linton (1975, 1978) and Wagenaar (1986) tested their memory for naturally occurring events, not artificial laboratory stimuli. For instance, Wagenaar recorded daily events in his own life for over six years, some 2400 separate events, and then tested his recall with combinations of four different cue types: what the event was, who was involved, and where and when it happened. Although he found that pleasant events were recalled better than unpleasant ones at shorter retention intervals, his evidence also showed that none of the events could truly be said to have been forgotten. Time-based cues, furthermore, were particularly useful in recalling events. In an interesting deviation from the original Ebbinghaus results (and from much laboratory research as well), Wagenaar also found that “the form of the retention curves cannot be explained by a supposed chronological ordering of events in memory, because incorrectly reproduced dates were equally often too recent and too remote” (p. 249). More powerful than the time lag, instead, was the salience of the event and the degree of emotional involvement.

The Relationship of Laboratory to “Real-World” Memory

I’d like to end this chapter with one final point and illustrate that point with one final experiment. You don’t have to agree with my opinion on this matter, of course, but thinking through your own position on the dispute should at least improve your understanding of the research that’s been described here (it may even provide new encodings of the material you’ve read, making that material more accessible to retrieval during your exams). The point is this. Autobiographical memory research tests retention for “real-world” information—names, faces, personal memories of past experiences, and so forth—rather than laboratory memory for the arbitrary and sometimes artificial lists of words that psychologists have been so fond of. The typical result is much better retention than we would have expected, based on laboratory demonstrations of forgetting or retrieval failure. Nevertheless, I believe it is a mistake to claim blithely that nothing of value has been discovered about human memory by means of laboratory experiments, to assert that the verbal learning tradition produced results as meaningless as the nonsense syllable was thought to be, or to say that we know little, if anything, about the operations of memory in ordinary life (as the quotation at the beginning of this chapter claimed). Instead, I would suggest that the recent research on autobiographical memory, important as it obviously is, would be impossible to conduct, and
The retention of information from the laboratory is strongly influenced by the conditions under which it is learned. When material is learned under conditions of high concentration and immediate recall, retention is much higher than when it is learned under conditions of low concentration and delayed recall.

The diagram above illustrates the relationship between the amount of material learned and the time intervals between learning and recall. The graph shows that retention decreases as the time interval increases, with a sharp drop in retention after the first few days. This phenomenon is known as the forgetting curve, and it is a common observation in the field of psychology.

The importance of early retrieval in the retention of information is emphasized by the fact that the retention of information is higher when it is retrieved soon after learning. This is due to the fact that retrieval strengthens the memory trace, making it more resistant to decay.

The forgetting curve demonstrates the importance of practice and review in the retention of information. The more frequently the material is reviewed, the better the retention. This is why it is recommended to review material regularly, especially in the early stages of learning.

In conclusion, the retention of information from the laboratory depends on the conditions under which it is learned, with high concentration and immediate recall leading to better retention. Early retrieval is crucial, and regular review strengthens the memory trace, making it more resistant to decay.

The forgetting curve highlights the importance of practice and review in the retention of information, emphasizing the need for frequent retrieval to maintain high levels of retention.
rehearsed less frequently. (A surprising result was that the roommates, who knew nothing of the experiment until shortly before testing, performed no worse than the subjects who had busily recorded all of the events.)

A skeptic might criticize the Thompson study by saying that the results were largely an artifact. That is, if you are able to recall a recent event easily, maybe this predisposes you to rate the event as having been more memorable originally, and to estimate that you had thought of it more frequently. In other words, maybe Thompson’s results were due entirely to a kind of circular logic—“If I can remember it now, it must have been memorable originally, and I must have thought about it frequently in the meantime.” How might Thompson argue against this criticism? It seems obvious that the retrospective ratings he obtained match literally hundreds of studies from the past, showing the effects of memorability, meaningfulness, and rehearsal on retention. It is precisely because his results conformed well to previous research, in which laboratory control permitted a careful evaluation of the variables, that we can accept those retrospective data as scientific evidence (in fairness, other internal consistencies within Thompson’s results support his conclusions as well).

Let me be clear here. I am not arguing for a return to the nonsense syllable and the P-A learning paradigm. As I indicated at the beginning of chapter 2, the topics being pursued by modern cognitive psychology seem inherently more interesting than those investigated some years ago—who wouldn’t rather study memory for meaningful material than the effect of nonsense syllable meaningfulness on P-A learning? And yet, bemoaning the “irrelevance” of that earlier work ignores the fact that it has contributed critically to our understanding of “real-world” memory. Our autobiographical memory research would have little idea of what variables might be important, either to manipulate or control, without the backlog of empirical research on laboratory learning and memory. And the tasks we use to assess memory for real-world experiences clearly rely on years of careful laboratory study and analysis (see, for example, Bahrick’s, 1979, “dropout technique” in the relearning task, motivated by years of empirical concern with degree of learning and overlearning in laboratory memory tasks).

To put it tritely, discarding our verbal learning tradition because of supposed “irrelevance” is “throwing the baby out with the bath water.” To be sure, cognitive psychology must concern itself with memory for real-world information, with the meaningful material that fills our heads and shapes our thoughts. But this doesn’t necessarily mean that we should turn our backs on the history and achievements of our discipline. Kintsch (1985) has stated this position succinctly, so we’ll rely on his expression of this sentiment:

[This] brings me to my final comment regarding the collective guilt feelings of Ebbinghaus’s successors about the ecological validity of their work. Ebbinghaus’s justification of the artificiality of his experimental procedures was a rea-

sonable one, for him. But 100 years later, we cannot continue to ask for more time and refuse to peer beyond the simplest laboratory phenomena. Fortunately, memory researchers today have no need to do so. The theory of memory that Ebbinghaus founded, and that reached its current form only in the last few decades, after Ebbinghaus’s theoretical limitations were finally overcome, does not need such excuses. True, it is still severely limited in scope, but the list-learning phenomena that have been studied since Ebbinghaus are reasonably well understood, and will form important building blocks of future, more comprehensive theories of memory and learning. True, there is much more to memory than what psychologists have chosen to investigate so far, but the basic mechanisms of encoding and retrieval that we have identified in our laboratories are of great help in understanding other cognitive phenomena in which memory plays a role, be it problem solving, discourse comprehension, or skill acquisition…  After the first 100 years, the discipline that Ebbinghaus founded is well on its way towards ecological validity (pp. 462–463).

Armed with our tools, recall and recognition tasks, reaction time and accuracy measures, and our explanatory principles, rehearsal, organization, imagery, interference, and the rest, let’s turn now to the other kind of long-term memory, semantic memory. You won’t be surprised, I suspect, when the first major topic we address, reconstructive memory, relies heavily on what you’ve just read—memory for real-world events and concepts is subject to, of all things, interference.

CHAPTER SUMMARY

1. The term metamemory refers to an awareness of how our memory system works and how to improve our learning and remembering. The classic method for improving such performance involves mnemonic devices, specialized rehearsal strategies that (a) ensure adequate storage of the information, and (b) provide a systematic method by which information can be retrieved. The classic mnemonic devices used a variety of techniques, especially visual imagery, to improve performance; familiarity with the mnemonic method provides a useful foundation for understanding both storage and retrieval effects in studies of memory performance.

2. Ebbinghaus was the first psychologist to conduct extensive investigations into the processes of learning and forgetting. Working on his own, Ebbinghaus invented scientifically acceptable methods of conducting such investigations; his use of the relearning task revealed a sensitivity to the demands of simple recall tasks, that such tasks tap consciously retrievable information, but may underestimate the amount of information actually learned and retained in memory. The classic forgetting curve he obtained, along with his results on practice effects, inspired the tradition of verbal learning and, indirectly, cognitive psychology.

3. “Storage” in long-term memory is somewhat of a misnomer, since subjects are often asked to “learn” words they already know, so storage, or transfer to long-term memory, means tagging or noting in some fashion
that a set of words was on the list to be learned. The important variables in storage are rehearsal and organization, whether these involve just verbal concepts or also include visual imagery and other nonverbal features.

4. Maintenance and elaborative rehearsal were generally viewed as having two distinctly different functions, the former for mere recycling of information without increasing its likelihood of retrieval, the latter for more semantically based elaboration and enrichment, where such rehearsal processed the information more deeply into the memory system and made it more memorable. Difficulties in this depth of processing framework involved definitions of the two types of rehearsal and specification of the notion of “depth.” When memory performance is tested with a recognition task, the results often seem to disconfirm the hypothesis that maintenance rehearsal merely maintains information, but does not make it more memorable.

5. Generally, the amount of rehearsal is positively related to recall accuracy for the primacy portion of a list. Organization, especially by category, but also by subjectively-defined chunks or clusters, improves memory performance because it (a) stores the information securely, and (b) provides a useful structure for successful retrieval. According to the encoding specificity principle, any related information that was encoded along with the studied information should serve as an effective retrieval cue. This is the case even if that related information is normally a lower-strength associate of the to-be-remembered material.

6. Interference was once thought to cause true forgetting from long-term memory. While interference is easily demonstrated in the laboratory using either proactive or retroactive interference tasks, the evidence now suggests that interference may disrupt retrieval. On this view, retrieval failure is not an issue of true forgetting, since information stored in long-term memory remains available in memory. Instead, retrieval failure is often due to loss of access to the stored information. Effective retrieval cues provide access to otherwise un retrievable information.

7. Studies of autobiographical memory, or memory in real-world settings, show the same kinds of effects as laboratory studies, but often more strongly. Recognition memory for information acquired across an extended period is remarkably accurate across many years, whereas recall performance begins to decline within months. Other autobiographical memory studies support the conclusions that long-term memory information is not truly forgotten, that memory for one’s own life events may be accessed especially well with time-based cues, and that distinctiveness or rated memorability is an important determinant of how accurately we can remember such events.

**SUGGESTED READINGS**

In addition to the papers cited in the metamemory section, there is an interesting commentary on the pros and cons of metamemory as a scientific construct at the beginning of the February issue of Child Develop-