Critical Concepts in the Study of Learning and Memory

Henry L. Roediger, III and Oyku Uner

Washington University in St. Louis


Corresponding author:

Henry L. Roediger, III
Department of Psychological and Brain Sciences, Box 1125
Washington University in St. Louis
One Brookings Drive
St. Louis, MO 63130-4899

e-mail: roediger@wustl.edu
fax: 314-935-7588
Abstract

Concepts are crucial in all scientific fields. They permit us to conceive of the phenomena studied in certain ways and, as the history of science teaches, they can prevent us from seeing these phenomena in other ways. The current chapter reviews and analyzes sixteen concepts critical to the study of learning and memory: Learning, plasticity, memory, encoding, consolidation, coding and representation, working memory, persistence (storage), retrieval, remembering, transfer, context, forgetting, inhibition, memory systems, and phylogeny and evolution. Of course, many other important concepts exist, but these are essential. Psychology and neuroscience seem to accrue new concepts and rarely is a concept abandoned.

Keywords: concept, learning, plasticity, memory, encoding, consolidation, coding and recoding, persistence, storage, working memory, retrieval, remembering, transfer, context, inhibition, forgetting
“With rare exceptions students of memory do not explicitly discuss the concepts they use in their work” Endel Tulving (2000, p. 34)

The empirical study of learning and memory is about 135 years old at the time of this writing, dating from Ebbinghaus’s (1885/1913) pioneering explorations. Even before the science of memory began, dozens of concepts referring to memory had been developed by philosophers, including association, forgetting, redintegration, and many others. In addition, some concepts about memory have been developed from words and figures of speech in everyday language—people often speak of searching their memories, and so do psychologists (Raaijmakers & Shiffrin, 1980).

Roediger (1979, 1980) argued that people have always had an implicit theory of mind and memory built into their language and thinking, in a way that is transparent and invisible. Most folk conceptions embrace the idea that memory is a large store or warehouse and memories are objects stored in this mind space; psychologists speak of short-term and long-term stores, adopting this implicit theory. Operations about memory in our conversations involve the same terms we would use for objects in an actual space. During learning, we speak of storing memories or acquiring them; of ideas sinking in or falling into place; we grasp ideas and they penetrate our minds, forming memories. We can organize our memories according to some plan (or schema), just as we can organize our possessions. And like spaces, minds can be narrow and shallow, or wide and broad. We keep or have memories in our mind space, and they may be in the top (or front) of our minds or in the back (or bottom) of our minds. Some memories may be concealed in the dark corners or dim recesses of mind. Operations for accessing memories are based on retrieving objects in our mental space. We speak of searching for or locating memories, or re-collecting them (collecting them again after we stored them). We can also lose a memory,
like we can lose an object, but we can suddenly find a memory, too. Watch for all these terms and more as you listen to people speak about their cognitive processes, and keep them in mind as you read this chapter. Some of our technical terms are borrowed from the vernacular.

This chapter is about more formal concepts, the ones scientists use, even if some of them are old ones that long predate scientific study. Concepts are critical to a field. They are the ideas that help us conceive of issues or problems in a certain way. Often science proceeds by scientists within a field discovering that the concepts they have been using no longer fit the facts and that a new conceptual framework needs to be created (e.g., scientists discarded the concept of phlogiston to explain fire). Thus lies scientific progress. Yet scientists often accept their concepts as self-evident truths; for example, of course people search their memories during retrieval. But is a search metaphor for retrieval really apt? Might some other idea be more appropriate (Wechsler, 1963)? Concepts are learned, accepted and used so readily that scientists rarely stop to scrutinize them deliberately, as we will do in this chapter (see Dudai, Roediger, & Tulving, 2007; Tulving, 2000).

Given the charge provided to us by the editors—to write a chapter on important concepts of learning and memory—one critical issue is how to even begin. After all, scientists from various backgrounds have developed hundreds of concepts in their study of learning and memory. Consider just some: Imagery, echoic memory, imprinting, synergistic ephory, habit, prospective memory, long-term potentiation, habit strength, and implicit memory. None of these concepts, important as some are, are included in our discussion below. Tulving (2007) asked, semi-seriously, “Are there 256 kinds of memory?” We can hope not, but in fact other candidates have been proposed since he published his list.
One issue that cuts across these concepts is that they differ in their level of analysis. Scientists approach learning and memory by focusing on the molecular and cellular components, or different parts of the nervous system, on networks of neural connections, on behavioral analyses of cognitive and comparative psychologists, and even on how learning and memory are affected by social context and social influences. The concepts below often pertain primarily to one or two levels of analysis. For example, neurobiologists study plasticity and consolidation, whereas for cognitive psychologists plasticity means learning and only recently has consolidation become a concept studied by behavioral methods. On the other hand, retrieval is a critical concept in cognitive psychology, but only rarely have neurobiologists considered it. These considerations of levels of analysis will occur repeatedly in this chapter. In fact, the same statement is true of this entire two-volume handbook.

How did we choose the most critical concepts for this chapter, as we cannot presume to cover all concepts in the field? We decided to rely on an earlier effort in which the first author took part. Along with Yadin Dudai, Susan Fitzpatrick, and Endel Tulving, the first author organized a conference around 16 critical concepts, and those concepts were chosen through many hours of face-to-face meetings, phone conversations, and e-mails. The conference was held in 2005 and a volume, Science of Memory: Concepts, was the result (Roediger, Dudai, & Fitzpatrick, 2007). The 16 concepts discussed briefly below are those selected for the earlier book, although of course our treatment of each here is brief. Whole books could be, and have been, written about some of them. We provide only a concise consideration of these concepts. Of course, we have updated some aspects of the concepts in light of new research since 2007.

Learning
Learning is a core concept for a field often referred to as “learning and memory.” Yet, surprisingly, it defies definition. As early as 1942, in his book *The Psychology of Human Learning*, John McGeoch argued that “The formulation of a systematic definition [of learning] is not easy, and the frequent attempts which have been made have not led to concordant results” (p. 3). He then defined learning operationally, from the perspective of a researcher working in the lineage of Ebbinghaus (1885/1913): “Learning, as we measure it, is a change in performance as a function of practice. In most cases, if not all, this change has a direction which satisfies the current motivating conditions of the individual” (1942, pp. 3-4). McGeoch went on to write an entire book on learning, definitional problems and all.

During the 20th century, many researchers worked on the problem of learning in the field of animal learning and behavior, studying both classical (Pavlovian) conditioning and instrumental (operant) conditioning. Again, they found it difficult to define learning. Hilgard (1948), in his classic *Theories of Learning*, wrote that “Learning is the process by which an activity originates or is changed through training procedures (whether in the laboratory or in the natural environment) as distinguished from changes by factors not attributable to training” (p. 4). He hastened to add: “The definition is unsatisfactorily evasive, and partly tautological…”. He went on to say that only “further discussion” could clarify matters. So, like McGeoch, he wrote an entire book about a concept that defies easy definition.

What are the problems? Consider just one: A person is given a task such as detecting one element of a display (say an X) in a field filled with other letters of the alphabet, screen after screen of them. The person might get faster over trials, which would be attributable to learning. But after hours on the task, performance would decline and that would be attributable to fatigue or boredom. Yet both changes would meet Hilgard’s first criterion for learning (and then be
tacitly excluded in the second part of his definition). Certainly, fatigue effects are not the only problem, as changes in motivation, maturation and other factors can affect performance on a task over time. Consider motivation: Suppose a rat that has learned to press a key whenever a green light comes on in an operant chamber is given ad lib food for an hour and then placed in the chamber. The animal no longer responds to the green light or does not do so as quickly and as vigorously. Has it forgotten the response? No, because if the animal is deprived of food for 24 hours, it will press the bar in the presence of the light. The motivation is simply lacking because the animal was sated. Such situations led to McGeoch’s qualifier about “motivating conditions” in his definition and “to changes not attributable to training” in Hilgard’s.

Rescorla (2007, p. 37) eschewed providing a definition of learning and he noted that “Contemporary textbooks in animal learning have by and large adopted the same view, typically simply sidestepping the issue of what learning is and introducing the subject by a kind of historical description.” Thus, many examples over the course of a textbook substitute for a formal definition that would attempt to encompass all types of learning. This pragmatic approach gets the job done, but it is still a bit unsatisfying because the concept is a key term in the field. Still, it prevents researchers from needless fights: We know learning when we see it.

Rescorla and Holland (1976) developed a framework for considering learning, preferring it to a formal definition. They suggested that all learning paradigms could be characterized by two points in time, t1 and t2. T1 refers to a state during which an organism interacts with the environment, whereas t2 is the later time where behaviors (or other changes, such as neural ones) in the organism are different as a result of the experiences. The difference between t2 and t1 reflects learning (with the same provisos as in the earlier definitions). This broad framework is handy, because it can encompass work of both researchers studying behavior of human or
infrahuman organisms, as well as neurobiologists studying neural changes as a function of experience. We will experience similar difficulties in defining memory, below. For now, we accept Rescorla and Holland’s (1976) broad view for purposes of this chapter. However, neurobiologists have preferred to cast learning with the term plasticity, to which we turn next.

**Plasticity**

Plasticity refers to change (or the capacity to change) that outlasts the cause triggering it (Moser, 2007). According to Konorski (1948), plasticity is one of the two meta-principles that contribute to the operation of the central nervous system. However, plasticity as a concept is too broad: Are we talking about plasticity at the synaptic level, behavioral level, or another level? Is plasticity adaptive? Does plasticity relate only to learning and memory? Is plasticity the neural underpinning of learning, so the same concept? We turn to these questions next.

In neurobiology, plasticity is commonly defined as structural changes at a cellular level that are necessary (and potentially sufficient) to generate durable and functional changes useful for the organism (De Zeeuw, 2007). Synaptic plasticity, in particular, has been widely studied to understand the molecular biology of learning and memory. However, focusing on one level of analysis can sometimes be problematic (Dudai, 2002). First, cellular plasticity does not only occur at the synapse. As an example, glia cells and intrinsic excitability of neurons also play a role in cellular plasticity (Allen & Barres, 2005; De Zeeuw & Yeo, 2005; Xu et al., 2005). Second, plasticity is a multi-level concept, with plasticity occurring at cellular, as well as network and behavioral levels (De Zeeuw, 2007; Moser, 2007). For instance, an area in the cortex allocated to represent sensory input can change based on the amount of received sensory input (Buonomano & Merzenich, 1998; Karni, Meyer, Jezzard, Adams, Turner, & Ungerleider, 1995). Finally, focusing only on cellular plasticity may create the assumption that it is necessary
and sufficient for learning and memory. Cellular plasticity is certainly necessary, but whether it is sufficient to result in learning and memory is unclear (Dudai, 2002; Martin, Grimwood, & Morris, 2000).

A critical issue regarding plasticity is whether changes are adaptive. Considering plasticity as an adaptive mechanism is tempting, because the concept is otherwise too broad. Moser (2007) argued that the immediate effects of drugs or the slow emerging effects of disease in the nervous system, for instance, would not be considered as plasticity by most neuroscientists. So, there seems to be an assumption that plasticity implies positive change. However, this approach is problematic for several reasons: 1) there is no a priori way to know what changes are adaptive, and 2) whether a certain change is adaptive can differ across contexts (Moser, 2007). If the definition of plasticity is to be constrained, perhaps it could be useful to count potentially adaptive changes as plasticity.

Of note, plasticity describes changes not just about learning and memory but also changes related to development, aging, and injury, to name a few. As Moser (2007) aptly pointed out, “plasticity is used to characterize a number of mechanisms at a number of analytical levels in widely different brain systems across all developmental stages” (p. 95). Plasticity is an umbrella term, but despite its breadth, it is an important principle of the nervous system.

**Memory**

As with learning, the other half of the phrase “learning and memory” is difficult to define precisely. The problem is that it is used in so many ways, and the main term is modified by numerous other adjectives (working memory, collective memory, long-term memory, semantic memory, eidetic memory, procedural memory, and so on; Dudai, 2002; Tulving, 2007). Because
so many varieties of memory exist, finding one broad definition to cover them all may be hopeless; if such a definition exists, it may be vacuous. Nonetheless, we persevere.

Here is a broad definition that Tulving (1983) suggested: “Memory has to do with the after-effects at one time that manifest themselves subsequently at a later time” (p. 7, cf., Rescorla and Holland’s (1976) definition of learning). We might consider the difference between performance at t2 relative to t1 as learning, with the aftereffect of the experiences considered memory. Tulving admitted that the definition might be too broad, but if one attempts to encompass the many expressions of memory, this is the type of definition that is needed. Of course, in this case it may be too broad, encompassing many phenomena that no one would classify as memory. If a player is hit by a squash ball today, tomorrow she may be sore and sporting a bruise. Thus, an event at t1 causes an after-effect at t2, but no one would refer to the bruise as a memory trace (although it would remind both players of what happened—a retrieval cue?). A definition needs to exclude cases as well as include them.

Schacter (2007) proposed four different senses of memory (crediting Tulving, 2000): “(1) The neurocognitive capacity to encode, store and retrieve information; (2) a hypothetical store in which information is held; (3) the information in that store; and (4) an individual’s phenomenal awareness of remembering something” (p. 25). The first of these definitions is most useful for researchers using behavioral or neuroscientific methods in studying learning and memory. The second has a more limited use (e.g., long-term store in some theories of memory). The third use could be the neural (e.g., cell assembly) or psychological (e.g., imagery) basis or code within the store. The last type, phenomenal awareness, is often associated with certain types of memory, but not all.
Many categories of phenomena that are often referred to as memory, or a type of memory, involve priming. Cofer (1960) seems to have first used the term priming, referring to the facilitation in behavior from a recent experience, but one that does not seem to (or need to) rely on conscious recollection. In the lexical decision task, the subject’s task is to react quickly to indicate whether or not a string of letters (chair or flirp) is a word. One of the first findings in using the task was that presenting a related word like table just before chair speeds the decision on chair by 40-50 msec (Meyer & Schvaneveldt, 1971). This is the phenomenon of semantic priming, and it occurs only across short intervals of time.

Yet priming of a different sort also occurs over much longer intervals. If chair is presented in a list and then subjects are asked, say, half an hour later to respond to word stems like cha____ with the first word that comes to mind, they are much more likely to say chair than any other word beginning with cha (charge, change, chart, charity, and others). In the control case, subjects are given cha____ without having studied chair, and they are less likely to respond with chair. Priming is defined as the difference between percentage completion with chair in the two cases (when primed by study and when not so primed). Notice that subjects are not told to remember words but just to produce the first item that comes to mind when seeing the word stem like cha____. Still, nothing in the procedure as described so far rules out using the cue to remember, either. However, various types of evidence do rule out that subjects use the cues to consciously remember when they are given instructions to respond with the first word they think of. For example, even densely amnesic patients show priming in this task, despite their ability for conscious recollection being greatly reduced by brain damage (Graf, Squire, & Mandler, 1984; Warrington & Weiskrantz, 1970). They can express recent experience in this indirect way by naming words, but not if they are directly asked to recall the information. Hence this form of
memory satisfies Schacter’s first characteristic above—encoding, storage, and retrieval of information—but not the others (especially the fourth one, phenomenal awareness). The terms used to describe many forms of priming are implicit tests of memory (Graf & Schacter, 1985; Schacter, 1987) or indirect tests of memory (Jacoby, 1984).

Admitting priming to the pantheon of “memory” opens up a Pandora’s box of other phenomena that meet the definition of a first experience changing behavior of a second. Take being vaccinated for measles as a child, an event that changes the immune system and has effects at a later time (being able to ward off infection if exposed to the measles bacterium). Is this memory? Scientists indeed refer to immunological memory, and the mechanism is what psychologists and neuroscientists would call priming. The first sentence of a paper in *Science* by Ahmed and Gray (1996) is illustrative: “The immune system can remember, sometimes for a lifetime, the identity of a pathogen” (p. 54). Psychologists would quibble with “remember” in that sentence, but immune memory would definitely classify as a form of priming (see Roediger, 2003, for other examples). And of course, memory is applied outside the body in other metaphorical ways, like the memory foam pillow that recovers its shape. Some plastics are said to have memory, too, because when deformed they can return to their original state when the right triggering condition is applied.

We have strayed too far afield. Let us be content with Schacter’s first option as our definition: Memory refers to the encoding, storage, and retrieval of information, with perhaps the proviso that the nervous system must be involved.

**Encoding**

Encoding refers to the initial registration or acquisition of information. Craik (2007) provided a more expansive definition of encoding: “the set of processes involved in transforming
external events and internal thoughts into both temporary and long-lasting representation” (p. 129). However, as with all our other concepts, both these definitions raise a whole host of questions that are difficult to answer precisely. For example, we conceive of the process of learning and memory as involving encoding (learning or acquisition), storage (or persistence of the memory trace), and retrieval (bringing the experience back to mind). Yet, for any event, when does encoding end and storage begin? This question is difficult, because encoding may extend over time, as in consolidation of the trace of the experience (see below).

To ensure that subjects encode memoranda, experimentalists may ask them to repeat aloud each word as it appears on a study list. Thus, accurate perception (encoding?) is guaranteed by one definition (accurate registration of information), but not by the definition of encoding as a process extended over time. Still, by one straightforward definition of encoding—accurate perception or acquisition—this technique suffices, as do others (e.g., asking subjects to make some sort of judgment on the words). However, people may rehearse (repeat subvocally) as they learn, which would extend the encoding process. And even when the event has left consciousness, consolidation may continue (as explained in the section below).

In many experiments, researchers manipulate variables during study, and often these are referred to as “encoding variables.” For example, Craik and Tulving (1975) developed an oft-used paradigm in which they manipulated the level of processing (Craik & Lockhart, 1972) during study or encoding. When presented with a word such as ORCHID, subjects would be asked either “Is it in upper case letters?” or “Does it rhyme with snorted?” or “Is it a flower?” To answer the first question, people only need to look at the case of the letters (visual features) of the word; to answer the second one they access the word’s sound (or phonetic) features; and to answer the third question they need to think of the meaning of the word to decide if it belongs in
the proper category. The answer in all cases is yes in this example, so the response is held constant. However, in most experiments, half the responses to the items are yes and half the time no. The finding from many experiments is that as the level of processing grows deeper, from visual to phonetic to meaning features, both recall and recognition improve. Hence, processing at encoding determined how information was recalled.

But did the effect only occur during encoding? Tulving (1979) argued that the effect may have occurred during retrieval. That is, perhaps semantic information is easier to retrieve than the other two types of information. Fisher and Craik (1977) showed that the levels of processing effect could be moderated by the type of retrieval cues used to probe retention, and Morris, Bransford, and Franks (1977) provided conditions in which phonetic encoding actually produced superior recall to semantic encoding when phonetic cues were used during the test. The overarching point is that differences observed in free recall or recognition as a result of “encoding manipulations” still need to take retrieval processes into account. No memory test measures what was encoded or stored; all answers are also filtered through processes operating during retrieval.

Another issue: Perhaps the course of consolidation differs after the different encoding manipulations. Should consolidation be considered part of the encoding process, or as a separate process that occurs after encoding? We turn to these issues in the next section.

**Consolidation**

How do we remember what we had for lunch yesterday, our home address, or how to ride a bike? Given how much we remember from our recent and distant past, some mechanism must exist that is responsible for converting our experiences into long-term memories. Consolidation is this process that enables what we encode to persist over time (McGaugh, 2000; Müller &
Pilzecker, 1900). According to the consolidation framework, our experiences leave traces that are initially labile and that need to be stabilized in order to be retrieved later. Consolidation explains the biological mechanisms with which newly acquired information remains available (Dudai, 2004; Glickman, 1961; LeDoux, 2007; McGaugh, 1966). However, similar to the other concepts we review, consolidation has its own definitional issues.

Two types of consolidation are believed to exist. Fast-acting consolidation at a local (i.e., synaptic) scale is called cellular consolidation, whereas consolidation at a broader (i.e., brain network) scale that takes much longer is called systems consolidation (Dudai, 2004; Kandel, Dudai, & Mayford, 2014, see also Chapter 6.7 in this volume). Cellular consolidation refers to the stabilization of synaptic and cellular changes produced as a result of acquisition (Dudai & Morris, 2000; Kandel et al., 2014). This type of consolidation occurs immediately after learning through a series of molecular processes. Drugs that inhibit protein synthesis, for example, prevent the conversion of short-term memories into long-term memories, providing evidence for consolidation that occurs locally (Davis & Squire, 1984; Dudai & Morris, 2000; McGaugh, 2000; Schafe & LeDoux, 2000). Systems consolidation, on the other hand, refers to the process through which memories that are initially dependent on the medial temporal lobe (or more specifically, the hippocampus and its surrounding areas) gradually become represented in the neocortex and no longer in the medial temporal lobe (Dudai, 2004; McClelland, McNaughton, & O’Reilly, 1995; Squire & Alvarez, 1995; Squire, Cohen, & Nadel, 2014). This process is considered to occur during offline periods like sleep and has a time course much longer than cellular consolidation (Diekelmann & Born, 2010; Dudai, 2004; Kandel et al., 2014). Currently, the existence of cellular consolidation is more widely accepted among researchers than the
existence of systems consolidation (for examples of views alternative to systems consolidation, see Nadel and Moscovitch, 1997, and Yonelinas, Ranganath, Ekstrom, and Wiltgen, 2019).

In the last two decades, however, researchers scrutinized the accuracy of the consolidation account (e.g., Nader, Schafe, & LeDoux, 2000; Sara, 2000). In particular, neurobiological findings showing that our memories go through a consolidation-like process during retrieval and behavioral findings showing that our memories are not always reliable challenged the views that consolidation happens only once after learning and that memories remain unchanged thereafter (Bartlett, 1932; Kida et al., 2002; Loftus, 1979; Misanin, Miller, & Lewis, 1968; Nader, 2003; Spear & Mueller, 1984; for other challenges to the consolidation account, see Nader et al., 2000). To address these findings, most researchers adopted the idea of reconsolidation that assumes consolidation is not a one-time event; instead, reactivation of memory traces makes those traces labile again and requires re-consolidation (Kida et al., 2002; Sara, 2000; Spear & Mueller, 1984). For instance, interfering with or blocking this process of reconsolidation can result in altered, weakened, or erased memories (Alberini, 2005; Nader, 2003; Sara, 2000). Of course, an important question is whether the original memory trace remains unchanged or if it is lost after several iterations of reconsolidation.

Originally, consolidation was an account of the biological changes that occur immediately after new learning and the changes related to that learning over time. The view that consolidation after acquisition stabilizes memories that are fixed is no longer endorsed. The current consolidation framework includes not just initial learning, but all reactivations of the memory trace resulting from that learning. Consolidation now refers to a more dynamic and ongoing process rather than a one-time event resulting in a fixed trace, as long as activation of memory traces occurs (Nadel, 2007; Sara, 2007). Irrespective of the changes in meaning,
consolidation is still a mechanistic explanation of how our experiences persist over time. But how is information represented so as to last? We turn to this issue next.

**Coding and Representation**

A key concept for cognitive psychologists is the idea of recoding, first introduced by George Miller in 1956. He pointed out that recoding of information is often one key to remembering well. The idea is that the environment might present information in one form, but that the cognitive system can convert, or recode, the information into a form more easily remembered. For example, a string of 15 digits—1, 4, 9, 1, 6, 2, 5, 3, 6, 4, 9, 6, 4, 8, 1—presented at a 1 sec rate would be exceedingly difficult for most people to remember in order. However, if a person were to recognize that these digits are the squares of the numbers from 1 to 9, they could recode them such that they would be easy to recall even a day later. This is a contrived example showing how a meaningless series of digits can be endowed with meaning to be learned easily, but the principle holds for all types of information. In fact, the basic tactic used by people competing in memory contests is to recode information given into a form that will make it easy to retrieve later (Foer, 2011).

Coding and representation mean different things to cognitive scientists and neuroscientists. Some neuroscientists consider how time, space, associations and more are coded at the level of synaptic connections, whereas other neuroscientists working with techniques like functional brain imaging seek networks of brain systems to understand coding. Unfortunately, these two approaches are not necessarily complementary. Treves (2007), writing from a more molecular perspective, argued that “the very principle of memory-based diversity in the messages carried by individual neurons in the same population provides an inviolable bound on the insight that can ever be obtained with imaging techniques such as functional magnetic
resonance imaging” (Treves, 2007, p. 58). That sentence is taken from an edited volume, and on the very next page, a different author begins his chapter on coding and representation with: “The study of human memory is at the edge of a new frontier thanks largely to functional neuroimaging” (McIntosh, 2007, p. 59). We may hope that both techniques can lend insight into issues of coding and representation.

Cognitive psychologists discuss coding in more abstract terms. To take one representative point of view, Paivio (1969) proposed a dual coding theory of how words, pictures and images are coded and stored, a theory that accounts for many findings. Words are represented by a verbal code, pictures by a nonverbal code. However, it is possible for words to evoke imagery (the word butterfly is likely to do so) and pictures can be named, invoking verbal codes. Paivio proposed that nonverbal codes are better remembered than verbal codes, and he and many others showed that pictures are better remembered than words, even on verbal recognition tests (e.g., Paivio & Csapo, 1969). That is, if pictures and words are presented in a list, and people are asked to recall the names of the pictures (i.e., words), recall of the studied pictures is better than recall of words. In addition, if subjects are instructed to listen to words and to recode them as mental images of the referents of the words (e.g., hear the word basketball and imagine a basketball), recall is better than if they recall words without such an instruction (Paivio & Csapo, 1969; 1973). The idea behind dual coding is that using both verbal and visual codes enhances recall or recognition over and above using either code, but especially the verbal code, alone. Paivio (1990) provided a book length exegesis of his theory.

Other cognitive psychologists have found it convenient to describe memory representations as composed of features (e.g., Bower, 1967; Underwood, 1969), but features are difficult to study empirically. Cognitive psychologists in the late 1960s and early 1970s tried to
empirically study representations in various ways, but these attempts had many problems and have largely been dropped from modern studies of memory (see Tulving & Bower, 1974, for a critical summary). Nonetheless, many contemporary models of learning and memory assume the existence of features of various types. The issue of memory representations is a critical one, and scientists are working to make progress on several fronts in neuroscience. Features may be conceived as attributes of an experience and now in the brain as attributes of neural activity. For example, in the past decade, a technique called representational similarity analysis has been applied to neuroimaging data to extract the features of encoded experience (see Rissman & Wagner, 2012).

**Working Memory**

The term working memory was first introduced by Miller, Galanter and Pribram (1960) to mean a short-term storage system. However, the term came into common use only after Baddeley and Hitch (1974) proposed a multicomponent model of working memory containing an executive capacity (central attention), a phonological loop (for producing and understanding spoken and written words), and a visuospatial sketchpad (to maintain nonverbal information). Baddeley (2000) added an episodic buffer to the model to integrate information from multiple sources and to provide a link to long-term memory.

The concept of working memory has changed over the years so that the central capacity aspect has received more attention. The idea today is that working memory is a capacity to hold information in mind briefly and work with it in the face of distraction. Working memory plays a crucial role in complex cognition, such as in mental arithmetic, text comprehension, and in future planning, among others. As an example, consider dividing a three-digit number by a one-digit number in your head (e.g., 138/3). You have to hold in mind the two numbers in the division, go
through each step, keep in mind each step’s product, update what you hold in mind as you go through the steps, and ignore any distracting information. A commonly used metaphor for working memory, mental workspace, successfully captures its key elements: temporarily storing information, processing that information, and avoiding distraction. That is, in the working memory system, people temporarily store newly acquired information as well as information from long-term memory, and manipulate them in order to execute a task at hand, while also avoiding distractors (Baddeley, 1992; Engle, 2002, 2018).

Working memory is characterized by its ephemerality; information from long-term memory and the current environment are brought to our focus only briefly. Because of its short-lived nature, working memory is related to short-term memory, but differs from it in one critical aspect: Whereas short-term memory refers to short-term storage of information, working memory is characterized by processing of that information, especially in the face of interference (Baddeley & Hitch, 1974; Cowan, 1988). In the division example above, short-term memory is recruited only to hold information in mind (e.g., what numbers you are dividing) without any manipulations to that information and without any distractions. Short-term memory can be conceived as one component of working memory.

In addition to independent stores (phonological and visuospatial), a critical component of the working memory system is the central executive that allows focused processing of temporarily available information (Baddeley, 1992; Baddeley & Hitch, 1974). Representations in working memory are transient; information can decay rapidly or can be subject to interference. So, the working memory system relies on active maintenance of these representations, which involves ensuring that the relevant representations are readily accessible and irrelevant.
representations are blocked or inhibited. As such, volitional attention plays a key role in the working memory system (Engle, 2002).

The working memory system is limited in capacity; however, arriving at an exact conceptualization of capacity is difficult due to its multifaceted nature. Some researchers have taken the approach to explain capacity by the number of items (or chunks) we can maintain in focus. For example, in his seminal paper, Miller (1956) argued that we can hold roughly 7 ± 2 items in short-term storage. Since Miller, others have argued that we can only hold about four chunks (Cowan, 2001; Watkins, 1974). However, the working memory system is not only limited by the storage mechanisms, but is also limited by executive attention, which is our ability to maintain focus on relevant information (Engle, 2002, 2018). As such, some researchers argue that working memory capacity is determined by how well we can maintain relevant representations in the face of distraction (Engle, 2002, 2018). In other words, executive attention is the predominant limiting factor of the working memory system rather than a precise number of items we can hold in mind (Engle, 2002, 2018; Engle & Kane, 2004).

Working memory underlies complex and flexible cognition, and its capacity strongly relates to a variety of constructs such as fluid intelligence and reading comprehension (Baddeley & Hitch, 1974; Daneman & Carpenter, 1980; Engle, 2001; Kane, Hambrick, & Conway, 2005). A system that relies on transient representations is, thus, essential to learning and memory. We now return to consider persistence of information.

**Persistence or Storage**

Persistence is the retention of what is learned over time. Similar to the other concepts we discuss in this chapter, there are key issues that should be considered when defining persistence.
What is the nature of the information retained over time? What are the challenges in studying persistence?

The metaphors used to describe memory, such as a library or a tape recorder, strongly imply that what is retained does not change over time, as in the early ideas about consolidation. Following this approach, the term persistence is sometimes referred to as *storage*. However, as discussed in the consolidation section, a wide variety of evidence exists illustrating the malleability of memory, suggesting that we do not always retain information in the same way it was acquired (e.g., Bartlett, 1932; Loftus, 1979). As an example, after people watch a car accident and are then asked to estimate the cars’ speed, the verb used in the question (e.g., smashed or bumped) can elicit different responses: People report higher speeds after the suggestion that the cars *smashed together* rather than *bumped into each other* (Loftus & Palmer, 1974). The term *storage* as well as the metaphors used to describe memory are, therefore, misleading. We consider persistence as a more appropriate term, although the encoding/storage/retrieval division serves as a convenient means of expression.

A critical issue about persistence is the difficulty in studying it. Psychologists cannot easily observe persistence without a behavioral expression, so studying persistence is often confounded with retrieval (but see neuroimaging studies that decode representations without or before overt retrieval responses, e.g., Norman et al., 2006). This is problematic because of two reasons. First, retrieval is not a passive event and it can alter our memory: Retrieval of previously learned information typically makes that information more accessible later (i.e., testing effect, see Roediger & Karpicke, 2006). In addition, retrieval can also inhibit recall of related information (i.e., retrieval-induced forgetting, see Anderson, Bjork, & Bjork, 1994), and it can make memories susceptible to alterations (e.g., Loftus & Palmer, 1974, and the research on
reconsolidation). Given that retrieval is not simply a window into what is retained, any study of persistence relying on overt responses will likely alter what is retained in some way. The reliance on retrieval to observe persistence is also problematic when we consider that not all retrieval failures illustrate failures of persistence. In some cases, even when information persists, we may not be able to remember it on one test (say, free recall) but we can on another type of test (cued recall or recognition). Tulving and Pearlstone (1966) distinguished between the availability and accessibility of memories. Psychologists want to know what information is “stored” in memory (what is available), but all that can ever be studied is what information is accessible under a certain set of retrieval conditions. The point is that the lack of behavioral expression is not enough to claim that information did not persist.

Howard Eichenbaum (2007) argued that persistence is critical in memory research, because there are no memories if neural representations and behavioral expressions do not persist (i.e., persistence is necessary). However, he also argued that persistence does not always lead to remembering (i.e., persistence is not sufficient). Persistence is an integral concept in memory, but only one part of the puzzle. It is also strongly linked to other concepts such as encoding and retrieval, which makes persistence difficult to isolate. Nonetheless, retention of previously acquired information is called persistence and the concept is critical in learning and memory research in conjunction with other concepts.

**Retrieval**

Tulving said during an interview that “The key process of memory is retrieval” (Gazzaniga, 1991, p. 91). This statement might seem hard to defend; after all, encoding and storage (or persistence) are also critical in the learning and memory process. However, we believe Tulving’s statement can be defended (Roediger, 2000). After all, encoding,
consolidation, and storage are cheap; every day these processes are carried out on the hundreds of events that every person experiences. Many of these daily events could be retrieved under free recall conditions a day later. Yet over time, the accessibility of these traces will vanish.

Returning to the distinction between availability and accessibility, many, and maybe even most, of the experiences of our day are encoded and stored; hence, they are available in the memory system. However, their accessibility depends on the retrieval cues in the environment at the time retrieval is attempted. Temporal, spatial, meaning, and other cues used in free recall might suffice for retrieval of events a day or two later; however, the events may not be remembered weeks later even with powerful retrieval cues.

The role of retrieval in the learning and memory process may be likened to the role of the perceiver in visual perception (Neisser, 1967). If a person walks into a busy scene (say, a professional sports stadium), the visual array of reflected light offers a huge number of opportunities for the perceiver. She must focus in on just some subset of light to be perceived and interpreted, with the rest passing by. Likewise, in memory, encoding and storage happen with nearly all events, leaving a huge array of potential experiences to be retrieved. Yet only a tiny portion of encoded and stored events will ever be retrieved and remembered, just as only a small segment of reflected light in a busy scene will be perceived.

Retrieval is governed by the encoding specificity principle (Tulving & Thomson, 1973; Tulving, 1983). The principle assumes that features of events are encoded and stored and that retrieval cues are effective to the extent that their encoded representation matches (or overlaps) the encoded trace. This principle can account for a large body of evidence from experimental designs in which conditions of encoding and of retrieval are varied so as to match or mismatch. One proviso on the principle is that cues should be specific enough to distinguish among traces
If a cue overlaps with too many traces, the cue is said to be “overloaded” and may not be effective. This issue concerning encoding/retrieval matches are considered more extensively in Roediger, Tekin, and Uner (2017).

**Remembering**

Like many concepts in the science of memory, remembering is a word used in English in many different senses. We can say “I remember that great movie I saw last week” or “I remember the major battles in World War II.” Yet the first event was personally experienced, and the second event was not (except for a few people still living). Tulving (1985) suggested a more specialized use for the term remembering, which has been adopted by many researchers. He proposed that the term *remembering* should be reserved for cases of personal recollection of events accompanied by retrieval of contextual details. Remembering the movie from last week would be a proper use of the term. On the other hand, he proposed that *knowing* should be the term used for retrieval from memory in cases where contextual details are absent or lost. Thus, *knowing* should be the term used in our recalling the major battles of World War II, because we were not there to experience them. However, we can also know events from our personal past. You may know you flew to Spain on vacation ten years ago, but remember nothing about the flight itself. As we will discuss in a later section, Tulving also proposed that remembering arises from the episodic memory system whereas knowing emerges via retrieval from semantic memory.

In addition to proposing the remembering/knowing distinction, Tulving (1985) provided a method to measure the two processes experimentally. The procedure is often used in recognition memory experiments; subjects are instructed that every time they recognize an event as old, they should next decide whether they remember the event from the list (recollect details
about the event’s presentation) or they just know it occurred in the list without remembering
details of presentation. Tulving (1985) also showed how the remember/know procedure could be
used for free recall and cued recall. Gardiner (1988; Gardiner & Java, 1990) was the first after
Tulving (1985) to use the remember/know procedure in analysis of experimental tasks. Over the
years, researchers have shown that variables can affect both remembering and knowing the same
way, or can affect remembering but not knowing, or can affect knowing but not remembering
(Rajaram, 1993).

A large literature on the remember/know distinction now exists, with some interpretations
different from Tulving’s, especially with regard to knowing (see Roediger, Rajaram, and Geraci,
2007, for an extended consideration among states of awareness during retrieval). Some
researchers believe that knowing arises from relatively automatic accrual of familiarity, so that
the distinction should be referred to as recollection and familiarity (e.g., Jacoby, 1991;
Yonelinas, 2002). Others have argued that the remember/know distinction reflects nothing but
differences in confidence; remembering arises when subjects are quite confident the event
occurred, and a know judgment is given when subjects are less confident (e.g., Dunn, 2004;
Rotello & Zeng, 2008). But this proposition may well have matters backwards—retrieval
experience is more likely to determine confidence than the other way around (Roediger & Tekin,
2020). In sum, Tulving’s (1985) distinction between remembering and knowing has turned out to
be quite useful to the field and is widely used in various forms today (see Umanath & Coane,
2020, for a recent review and critique).

**Transfer**

Transfer is defined as the effect of practicing one task on learning or performance of
another task. Transfer is so ubiquitous that coming up with a more specific definition is difficult.
In an attempt to capture its complexity, Barnett and Ceci (2002) provided a taxonomy of transfer with several dimensions that comprise different types of transfer: the kind of skill transferred, how transfer is measured, the memory demands of the task, as well as the similarity of the two tasks regarding modality (auditory, visual and others), context (physical, temporal, functional, and social), and knowledge domain. Using knowledge about changing a flat tire on a bicycle to change a flat tire on a car is an example of transfer, and so is identifying the artist of a new painting after having studied the artist’s paintings. These are both instances of positive transfer; that is, previous experiences enhance performance on another task. Negative transfer occurs if prior experiences impair performance on a second task, such as after studying chemistry when a student misidentifies a chemistry problem and applies an incorrect formula.

Early investigations of transfer were focused on the idea of formal discipline in education. The goal of formal discipline was to train students on one task or subject in hopes of providing a general ability that would transfer to learning in other contexts (e.g., memorizing poetry should generally benefit one’s memory, and thus learning history would be improved). However, researchers have consistently failed to show evidence for this kind of transfer, beginning with Thorndike and Woodworth (1901). Current variants of this pursuit are brain training games or working memory training, both of which are expected to improve performance on highly dissimilar tasks such as those needed in daily life. However, in controlled experiments, results show that learners improve on the training task and sometimes on highly similar tasks, but the skills they learn do not transfer to dissimilar tasks, which is the intent of the training (Shipstead, Redick, & Engle, 2012; Simons et al., 2016). In other words, such cognitive (or brain) training may transfer to highly similar tasks (called near transfer) but not to dissimilar
tasks (far transfer). Obtaining far transfer has proved difficult in many domains, such as education, not just in cognitive training.

Despite the disappointing lack of transfer observed in some cases, transfer effects are more common, though not always present, in domains such as skill learning, problem-solving, and concept learning (Ashby & Maddox, 2005; Gick & Holyoak, 1980, 1983; Goldstone, Kersten, & Carvalho, 2017). The primary objective of skill learning is to ensure learners are able to use these acquired skills in new situations. For example, someone who learns to play the piano should be better able to learn a new piece of music or play in a different venue from the one in which piano practice occurred. Transfer, in this context, is a measure of expertise of that skill. Similarly, the goal of problem-solving research is for learners to be able to use previously learned problem solutions in novel instantiations of problems. A challenge in this kind of problem-solving is recognizing that the new problem is similar to a previously studied problem despite differences in surface features (Gick & Holyoak, 1980, 1983). Without noticing that a given problem is akin to a known problem, learners will likely fail at transfer. Fortunately, in the concept learning domain, a major focus has been understanding how learners identify novel instances of a concept after having seen several of its instances (Ashby & Maddox, 2005; Goldstone et al., 2017). Alternating exemplars of different categories, for example, usually leads to better classification of new exemplars from those categories than studying all examples of one concept and then all examples of a different concept (for a review, see Carvalho & Goldstone, 2019). Findings from these domains (skill learning, problem-solving, concept learning) have important educational implications: Students need to learn new languages, solve complex mathematics problems, understand great works of literature and philosophy, to name a few.
Transfer, therefore, is a necessary component of nearly every students’ understanding of course material.

Why is transfer hard to observe in some situations (such as memorizing poetry to improve memory of history), yet is more common in others (such as identifying the species of a new bird when one has knowledge of birds)? Similarity of learning and transfer contexts is a key factor in determining whether transfer is successful. As noted above, the terms near transfer and far transfer describe how similar these two contexts are, though of course similarity is a continuous dimension (Barnett & Ceci, 2002). Similarity can also be determined by the processes one is engaged in during learning and transfer. If the processes required by the learning and transfer tasks are similar, the influence of the former task on the latter will be more readily apparent, a finding known as transfer-appropriate processing (Morris, et al., 1977; see also Kolers & Roediger, 1984).

Although the definition of transfer needs further specifications in any particular instance, its general definition as the influence of previous experiences with a task on learning or performance of a different task encompasses the fundamental idea.

**Context**

Context is another broad concept. At the most general level, it refers to the situation or circumstances in which an event takes place during learning and testing. The term may refer to external context, such as a conditioned stimulus or the setting in which learning takes place, or it may refer to internal context. Internal context may refer to semantic context (induced by recent words seen, for example), or it may refer to situations such as the mood state (happy or sad) or drug state (inebriated or sober). In conditioning experiments, the context is often conceived as the stimulus situation such as a light, a tone, or even the features of the conditioning chamber.
We will focus here on context in human learning and memory experiments, but we cannot cover all types of context.

Context can be described as either local (specific) or global (general), but again this dimension represents a continuum and can depend on how rapidly contextual features diminish over time. In paired associate learning, learning *giraffe-dewdrop* so that on a later test *giraffe* is given as a cue (to recall *dewdrop*), *giraffe* would be the local or specific context for the response. The language of conditioning experiments with infrahuman animals is often carried over to discuss paired-associate learning, so the left-hand member is referred to as the stimulus and the right-hand member the response. Since the 1970s, the language has changed such that the left-hand member is often referred to as the context during encoding and the cue during retrieval, with the right-hand member referred to as the target. Often the phrase cue-target pair is used, even though, properly speaking, the word cue refers to retrieval cue used at testing and not the context in learning.

A large literature exists on cue-target relationships in paired-associate learning, both in the older S-R tradition and in the newer tradition that began with studies on encoding specificity in recall and recognition. For example, one dramatic finding was that in learning pairs like *night-day*, *night* can be a more effective retrieval cue for *day* on a recognition test than *day* is for itself (Tulving & Thomson, 1973). This phenomenon of recognition failure of recallable words provided compelling support for the encoding specificity principle, and showed that the common-sense claim that recognition is always superior to recall is not true. One interpretation is in terms of semantic specificity: The meaning of *day* is different in the context of *night* than when *day* is seen by itself, and this variation changes the effectiveness of the cues for the target.
At the global level of context, researchers have studied the effects of drugs on memory. Almost all drugs that depress the central nervous system impair learning and memory. That is, if the drug has been ingested or injected before learning, later retention when the drug has worn off will be impaired relative to a condition in which no drug was administered during learning or testing. The amnestic qualities of depressive drugs are well known. However, researchers later discovered the phenomenon of state-dependent retrieval. If people are given the drug, say marijuana, both before learning and before testing, recall of information is better than in the condition when the drug is given before learning but not at testing. That is, if people are mildly intoxicated when learning, they remember better if they are intoxicated again during testing than if they are sober during testing (e.g., Eich, Weingartner, Stillman & Gillin, 1975). This phenomenon is referred to as state-dependent retrieval. Eich (1980) showed that the effect is much more likely to occur on tests of free recall than on recognition or cued recall. In these latter cases, the specific, focal retrieval cues seem to override the more global contextual cues.

Global physical context has been postulated to operate in a similar manner. That is, when people have experiences in one situation, they are more likely to retrieve the information if they are placed back in that same context than in a different context. Smith, Glenberg, and Bjork (1978) reported several experiments that revealed large differences in free recall when subjects learned lists in one distinctive type of room and then recalled the lists in either the same room rather than a quite different room. However, Fernandez and Glenberg (1985) failed to replicate this outcome even when they used the same rooms as Smith et al. (1978). One possibility is that, for unknown reasons, students in Smith et al.’s experiments may have integrated words with the room environment more than in the replication attempts, although of course there is no compelling reason to suspect this possibility. Still, others have obtained the room-context effect
(for a meta-analytic review, see Smith & Vela, 2001), although the effect sizes are quite modest. Eich (1985) obtained a room-context effect when he instructed subjects to integrate images of words in a list with objects in the room, but not when subjects imagined the objects floating in space. Smith (2007, p. 111) suggested that “Contexts can vary in terms of how enmeshed they are with the focal events for which they serve as context.” The more enmeshed the objects are, the greater the possibility of room-context effects. In more recent research, Smith and Manzano (2010) showed much more powerful effects when using video recordings as context, with words superimposed over them. The words were much more likely to be enmeshed with the video than when they are studied in rooms with the room serving as a static background (see Smith, 2014, for a review).

We have not taken the space here to describe other important types of context, such as the temporal context of events and the role context reinstatement in episodic memory processes (e.g., Howard & Kahana, 2002). Suffice it to say that context is a critical concept in understanding memory, although a concept instantiated in many ways in practice (see Chapter 5.12 for a further discussion of context and context reinstatement effects).

**Forgetting**

Forgetting is a common problem that afflicts all humans. We all have failed to retrieve information when we want to, be it a movie, a name, a restaurant, or in response to a question on a test. As with other phenomena, there is both a general meaning of forgetting that everyone has as well as more technical definitions provided by psychologists. Here is one: Tulving defined forgetting as “the inability to recall something now that could be recalled on an earlier occasion” (1974, p. 74). This definition implies that information was encoded, stored, and even retrievable, but then forgotten after that time.
One strong form of the term forgetting is the complete erasure of the memory trace of an experience from the nervous system. Davis (2007) considered this meaning as well as possible evidence for it: “Only when all the cellular and molecular events that occur when a memory is formed return to their original state would I say this would be evidence for true forgetting” (Davis, 2007, p. 317). The evidence required by Davis could only be found in simple nervous systems (e.g., in gastropods or slugs) in which the entire neural circuitry has been mapped out, and we are unaware of any evidence in the intervening years that would satisfy his criterion. Does that mean that “true forgetting” never occurs? We think it probably does. After all, we encode and store hundreds, maybe thousands, of events each day, mostly small ones. We could recall some of them in an hour and some the next day, but in 30 years we would have no chance in recalling or recognizing what we did on that particular date. Of course, it is possible that what you had for dinner on April 22, 2005 is stored in your brain (even if not retrievable), but equally likely is that it has been erased in Davis’s sense or has at least been overwritten by so many similar events. McGeoch (1932, 1942 Chapters VIII and IX) argued that forgetting was caused by several factors, which included a change in context between learning and testing, which means the cues at the time of testing would not match those of learning.

Tulving championed the idea of forgetting as retrieval failure (e.g., Tulving & Pearlstone, 1966; Tulving & Psotka, 1971). He and his colleagues showed that forgetting on a free recall test could be reduced (although not completely eliminated) by providing powerful retrieval cues, such as category names for words belonging to the categories that had been studied. Many other demonstrations of reversals of forgetting by providing cues exist (for a review, see Roediger et al., 2017). Forgetting as retrieval failure is accepted by many psychologists (Loftus & Loftus,
1980), but of course one can never completely rule out the possibility that some forgetting is permanent.

A third definition of forgetting is the loss of information over time. For example, seven different groups of subjects may learn some material to the criterion of, say, 80% correct. Then one group is tested five minutes later, another group an hour later, a third group five hours later, and so on. The test could be recall, recognition, relearning of the material, as well as other measures. Ebbinghaus (1885) was the first person to report such an experiment, although in his case there was only one subject (Ebbinghaus) and he learned and tested himself over different sets of material for different periods of time. In nearly all cases of such experiments, Ebbinghaus’s basic result is replicated: Forgetting occurs quickly over an early period of time and then levels off. Ebbinghaus proposed that the forgetting curve was logarithmic, and in reviewing 100 years of forgetting, Rubin and Wenzel (1996) concluded he was right. However, Wixted and Carpenter (2007) provided analyses and evidence that a power function fits forgetting curves slightly better than a logarithmic function.

Why does forgetting occur? Several theories exist to explain forgetting, but we cannot take the space to review them here. Ninety years ago, McGeoch (1932) proposed three factors that cause forgetting: competition among memories (referred to as cue overload in the previous section), contextual change, and a change in retrieval sets or cues. All three factors are still used in current theories of forgetting. Roediger, Weinstein, and Agarwal (2010) provide a brief review. Their chapter appears in an entire book on forgetting by Della Sala (2010), which provides an excellent resource.

**Inhibition**
The concept of inhibition implies a comparison: One condition or state is inhibited relative to another condition or state. For example, lateral inhibition in the nervous system refers to the effect that an excited neuron has on its neighbors, because the excited neuron slows the firing rate of the neighbors. That is, the neighbors will fire more if the neuron is not excited (the comparison) than if it is. Neuroscientists study inhibitory phenomena and their interaction with excitatory phenomena, which can lead to complex and nonlinear patterns of neural activity (Buzsaki, 2007).

Inhibition as a general concept has been imported into the study of psychological processes, but often the underlying neuroscientific understanding of these inhibitory phenomena is lacking. In Pavlovian procedures, conditioned inhibition refers to the case where a conditioned stimulus (CS) becomes a signal for the absence of the unconditioned stimulus. In this condition, the underlying behavior (the response) is inhibited relative to the case when the CS is absent (Rescorla, 1969).

In the study of human learning and memory, inhibition is used in several different ways. For example, Slamecka (1968, 1969) showed that after subjects have studied a list of words, giving them part of the list as cues to help recall the remainder of the list actually hindered them in recalling the other list items. This part-list cueing inhibition is relative to free recall with no cues; thus, giving part of a list as cues inhibits recall of the remainder of the list.

Another type of inhibitory phenomenon is called retrieval-induced forgetting (Anderson, et al., 1994). In this case, conditions are arranged such that subjects are induced to retrieve some items from a set before they recall the remaining items. Recall of the remaining items is impaired (inhibited) relative to the case where prior recall does not occur (see too Roediger, 1978). In this case, inhibition is also used to describe the mechanism underlying the phenomenon in addition to
the phenomenon itself. That is, the act of recall is said to inhibit (weaken) the representation of other items and hence induce their forgetting. Other explanations of retrieval-induced forgetting are possible (Raaijmakers & Jakab, 2013).

Another case in cognitive psychology where inhibition is used as an explanatory concept rather than to describe a phenomenon is in theories of retroactive interference. For example, if one group of people learns a list of pairs of A-B words like *cat-orange* and *book-chimney*, they will remember these pairs less well later if they also have to learn a second list of A-D pairs like *cat-peacock* and *book-heaven* before they are tested. The control condition in this case learns the A-B list and then either performs an unrelated filler task or learns a C-D list with no words in common with the first list (this control provides nonspecific interference). On the final test, when given cues like *cat* and *book* and asked to recall the word from the first list, subjects perform more poorly in the A-B, A-D condition than in the conditions in which subjects learned the C-D second list or performed an unrelated task. This phenomenon is called retroactive interference and it is thought to be a major cause of forgetting. One factor thought to underlie retroactive interference is response inhibition. That is, when given the A cue like *cat*, the responses *orange* and *peacock* compete for conscious access, and recall of the correct response (*orange*) is thought to be impaired or inhibited relative to the control conditions by subsequent learning and retention of cat-peacock.

In sum, inhibitory phenomena are important and much studied in the science of learning and memory. Sometimes the term inhibition is used to refer to a phenomenon (part-list cueing inhibition) and sometimes the term is used to refer to the mechanism underlying an interference phenomenon (see Chapter 6.3 on Inhibition Theory).
The general set of human abilities referred to as “memory” can be subdivided into different types with partially independent neural systems underlying the types. Schacter and Tulving (1994, p. 13) discussed several characteristics of memory systems: 1. There are different cognitive and behavioral functions of the systems and different kinds of information being processed; 2. The processes operate according to different laws or principles; 3. The systems have different neural substrates; 4. The systems have different times of appearance in ontogenetic and phylogenetic development; and 5. They have different formats of information represented in the system.

Here we will briefly discuss five systems that many researchers agree upon: Episodic memory, semantic memory, the perceptual representation system, primary or working memory, and procedural memory (Schacter & Tulving, 1994). We have discussed working memory in an earlier section, so here we will briefly describe the other four.

Tulving (1972) proposed the basic distinction between episodic and semantic memory, but the concept of episodic memory has evolved over the years. Briefly, episodic memory refers to our recollection of personal happenings dated in a particular time and particular place, often with rich details. Episodic memory is the only memory system dedicated to the past, but it also enables people to think about the future in concrete terms. Episodic memory records the events of our lives and is usually what laypeople think of as “memory.” When amnesia occurs due to an injury to the brain or from a disease like Alzheimer’s disease, the memory system lost is episodic memory. The other memory systems largely remain intact, at least in early stages of the disease.

Semantic memory refers to our atemporal knowledge of the world—that Rome is the capital of Italy, that dogs have four legs and bark, that the chemical symbol NACL stands for salt. Semantic memory is our storehouse of knowledge, of word meanings and of any other sort
of meaningful knowledge. People encode events in episodic memory by using information in semantic memory.

The perceptual representation system is concerned with identifying perceptual objects in the world, including words (Tulving & Schacter, 1990). This system is responsible for phenomena of perceptual priming, such as its specificity of modality. For example, seeing words during a study phase produces more priming on a visual test than hearing the words during study (e.g., Rajaram & Roediger, 1993) whereas the reverse is true in auditory priming tasks (Pilotti, Gallo & Roediger, 2000).

Procedural memory is the system that underlies motor performance or skills. Tying one’s shoes, serving a tennis ball, riding a bicycle are all examples. Procedural memory refers to a “vast category, as yet largely unexplored and unknown” with “a large number of rather specific subsystems” (Schacter & Tulving, 1994, p. 16).

Squire (1987) proposed a hierarchical arrangement of memory systems, with the two main divisions being between declarative memory (knowing that something occurred) and procedural memory (knowing how to do something). The distinction between knowing that and knowing how originated in writings by Gilbert Ryle (1949). In Squire’s framework, episodic and semantic memory are subsystems within declarative memory. Procedural memory is thought to have several specific subsystems, too. Squire’s theory has also developed and expanded over the years.

We are sketching these memory systems with broad strokes and much too briefly. Whole books have been devoted to various systems (e.g., Tulving, 1983). Further, as noted, ideas about these systems have developed and changed over the years, although they retain their fundamental distinctions.
Debate about the number and arrangement of memory systems in the brain proceeded from the 1980s through perhaps 2010, but recently research has turned to brain networks uncovered by functional neuroimaging (e.g., the parietal memory network; Gilmore, Nelson & McDermott, 2015; McDermott et al., 2017). The various networks discovered by fMRI do not map neatly on to proposed memory systems, so the network approach may offer an alternative way of conceiving of neurocognitive systems for memory.

**Phylogeny and Evolution**

Unlike other concepts, phylogeny and evolution are not specific to learning and memory. All biological systems have evolved over time. Still, these concepts are critical for understanding learning and memory. Because all organisms have the capacity to learn, the issue of how learning and memory differ across species is critical. Do some basic principles and mechanisms of learning seen in simple organisms like *Aplysia* (snails) or *Drosophila* (fruit flies) underlie learning in species with more complex nervous systems, including the human nervous system? Many neurobiologists who study learning assume this to be true.

One method of addressing questions of phylogeny lies in comparison. “Comparison is the magic tool of evolutionary biology,” according to Menzel (2007, p. 371). Understanding how learning varies across species leads to studies in comparative psychology. One tactic is developing tasks that aim to measure the same ability as a function of the same independent variables across species. Pavlovian fear conditioning and operant conditioning also occur in most species and probably in all vertebrate species. As a more specific example, Wright (1998) developed procedures that permitted serial position curves to be created from visual and auditory lists of elements by using a recognition memory paradigm that was generally the same for three species (pigeons, monkeys, and humans). He showed that serial position curves for auditory and
visual presentation of items differed and that these differences changed over time. Further, the pattern of results was the same for all three species, although the time courses differed. The assumption, then, is that similar processes underlie serial position functions and for both visual and auditory presentation of items and how they change over time.

As species evolved more complex nervous systems, we assume that learning and memory abilities evolved, too, and became more complex. For example, Tulving (2005) argued that episodic memory, marked by specificity of time and place of a memory, is uniquely human. However, Clayton and Dickinson (1998) have argued that scrub jays (a member of the crow family) have a remarkable ability to “remember” where, when, and what type of food they stored previously. They argue that the birds’ abilities at least mimic those of humans, and they refer to episodic-like abilities in the scrub jay. The scrub jay example may represent an example of convergent evolution; that is, the ability to use time and place of occurrence in memory may have evolved separately in scrub jays and humans rather than having been inherited from a common ancestor (which would represent homology).

Testing evolutionary ideas about learning and memory is difficult. However, it is possible. For example, Pravosudov and Clayton (2002) studied black-capped chickadees from either Colorado or Alaska. Because winters in Alaska are much more severe and longer than those in Colorado, the authors argued that selection pressure should have led the Alaskan chickadees to have developed more accurate memories for where their food caches were placed. When they tested both types of birds in identical laboratory environments, they found evidence supporting the prediction.

Can evolutionary ideas about remembering be tested in humans? James Nairne has developed a survival processing paradigm that is intended to do just that, and he created it on the
territory of experimental psychologists’ bedrock technique for studying memory: Recalling lists of words. Now, no one believes that human memory evolved to remember lists of words, but Nairne, Thompson, and Pandeirada (2007) developed instructions given before learning a list of words intended to evoke survival processing in people. They gave students detailed instructions asking them to consider how words in a list would help them to survive in a grassy plain (like the ones in which humans are assumed to have evolved). They compared this instruction to other instructions known to aid learning (e.g., instructions for rating words on pleasantness in one condition, or rating how much the words are relevant to the person in another condition). They found that learning words while thinking about survival caused them to be remembered better than under the other instructional sets, and they replicated this finding several times. In addition, many other investigators have replicated the basic effect, although some have attempted to explain it in terms of already-known memory mechanisms (e.g., Burns, Hart, Griffith & Burns, 2013). Discovery of the survival processing effect began an approach and a program of research known as adaptive memory that has continued to develop (see Nairne, Pandeirada, & Fernandes, 2017, for a review).

Of course, other approaches to adaptive memory exist. Klein (2007) outlines another, complementary, approach that argues that the concept of the self is critical in understanding memory in general and episodic memory in particular. The study of evolutionary approaches to learning and memory, particularly in humans, remains in its infancy. The topic is critically important and we await future discoveries and theories.

**Conclusion**

After covering sixteen concepts, which may seem like too many to our exhausted reader, we cannot help but think that we have only scratched the surface. So many more important
concepts exist. We are tempted to start listing another sixteen, but even that would not cover the field. We can only ask readers to keep their eyes open for other concepts, asking if they are needed to our field. Are there any we can jettison? Has our field reached the mature state, as with the example of chemists and phlogiston, where we can say, “We used to think that was a critical concept, but now we know better”? The authors cannot think of a single case where this has happened. Aristotle believed that associations were important for memory and they were formed by contiguity of elements. Neuroscientists and psychologists still believe that today, 2300 years later, though we know much more about conditions needed for associations to form.
References


Carvalho, P. F., & Goldstone, R. L. (2019). When does interleaving practice improve learning?.


Dudai, Y. (2002). The neurobiology of consolidations, or, how stable is the engram?. *Annual Review of Psychology, 55*, 51-86.


Tulving, E. (1974). Cue-dependent forgetting: When we forget something we once knew, it does not necessarily mean that the memory trace has been lost; it may only be inaccessible. *American Scientist, 62*(1), 74-82.


