VERBAL LEARNING AND MEMORY:
Does the Modal Model Still Work?

Alice F. Healy and Danielle S. McNamara

Department of Psychology, Muenzinger Building, University of Colorado, Campus Box 345, Boulder, Colorado 80309-0345

KEY WORDS: memory models, primary memory, sensory memory, secondary memory, working memory

ABSTRACT
This chapter focuses on recent research concerning verbal learning and memory. A prominent guiding framework for research on this topic over the past three decades has been the modal model of memory, which postulates distinct sensory, primary, and secondary memory stores. Although this model continues to be popular, it has fostered much debate concerning its validity and specifically the need for its three separate memory stores. The chapter reviews research supporting and research contradicting the modal model, as well as alternative modern frameworks. Extensions of the modal model are discussed, including the search of associative memory model, the perturbation model, precategorical acoustic store, and permanent store. Alternative approaches are discussed including working memory, conceptual short-term memory, long-term working memory, short-term activation and attention, processing streams, the feature model, distinctiveness, and procedural reinstatement.

CONTENTS
INTRODUCTION ................................................................. 144
BACKGROUND ..................................................................... 144
PRIMARY MEMORY ........................................................... 146
Modern Extensions of the Modal Model ................................ 148
Additional Memory Systems ............................................. 153
Unitary Memory System .................................................. 159
Summary ........................................................................... 161
SENSORY MEMORY .......................................................... 161
Precategorical Acoustic Store ........................................... 161
INTRODUCTION

This chapter focuses on recent research concerning verbal learning and memory. A prominent guiding framework for research on this topic over the past three decades has been the modal model of memory (Atkinson & Shiffrin 1968, Glanzer & Cunitz 1966, Waugh & Norman 1965), which postulates distinct sensory, primary, and secondary memory stores. Although this model continues to be popular, it has fostered much debate concerning its validity and, specifically, the need for its three separate memory stores. In this chapter, we review research supporting and research contradicting the modal model as well as alternative modern frameworks. We begin by summarizing the empirical support for the original model. The remainder of the chapter addresses in turn issues concerning each of the three memory stores. 1

BACKGROUND

An initial statement of what has since been termed the modal model can be traced to James (1890), who distinguished between primary and secondary memory. James described primary memory as that which is held momentarily in consciousness and secondary memory as unconscious but permanent. An important impetus to modern versions of the modal model was the discovery that a short sequence of items is forgotten within seconds when rehearsal is prevented by a distractor task interpolated between item presentation and recall (Brown 1958, Peterson & Peterson 1959; but see Melton 1963). Other findings that also invited the distinction between primary and secondary memory included neuropsychological studies of amnesic patients unable to form new long-term memories (Milner 1966, but see Graf et al 1984) and studies showing that short-term memory (STM) tended to rely on phonetic coding and long-term memory (LTM) on semantic coding (Baddeley 1966, but see Shulman 1971).

1 It should be noted that this chapter is not intended to cover the entire spectrum of recent research on verbal learning and memory. For example, space limitations do not allow us to review indirect measures (Richardson-Klavehn & Bjork 1988), mathematical models (Raaijmakers & Shiffrin 1992), or neuropsychological studies (Squire et al 1993) of memory.
Modern descriptions of the modal model were presented by Waugh & Norman (1965) and Glanzer & Cunitz (1966), and the fullest description was provided by Atkinson & Shiffrin (1968), who added sensory memory to the primary and secondary memory dichotomy. Specifically, Atkinson & Shiffrin postulated three distinct memory stores: sensory registers (with separate registers for different senses, including visual, auditory, and haptic), short-term store (STS or primary memory), and long-term store (LTS or secondary memory). In terms of a computer metaphor, these stores constitute the essential permanent structural features, or hardware, of the system. In addition, Atkinson & Shiffrin postulated various control processes, or subject strategies, constituting the software of the system. One control process was given most extensive consideration: rote rehearsal. The concept of a rehearsal buffer in STS was used to describe this process.

Although much of their empirical work involved the continuous paired-associate paradigm, Atkinson & Shiffrin (1968) also devoted considerable attention to performance in the standard free-recall task. According to the simple version of their model applied to free recall, each stimulus item enters a fixed-capacity rehearsal buffer and displaces a randomly selected item already there when the capacity (about four items) is exceeded. As long as an item is in the buffer, information about it is transferred to a permanent LTS. The amount of information transferred is a linear function of the time in the buffer. At the time of test, subjects initially output the items still remaining in the buffer and then make a fixed number of searches of LTS. Crucial to this version of the modal model, as well as to earlier versions, are the assumptions that an item may be retained in the STS buffer as well as in LTS at the same time and that recall of any particular item, including those presented most recently, can derive from information in both STS and LTS.

The Atkinson & Shiffrin (1968) model accounts for the bowed serial position function and the effects of such variables as the presence of a distractor task, the rate of item presentation, and the list length. Specifically, the advantage for the most recently presented items (the recency effect) is explained by the fact that those items remain in the buffer at the time of test. The advantage for the items presented initially (the primacy effect) is explained by the fact that those items stay in the buffer longer than subsequent items. Because the buffer starts out empty, the initial items are not displaced by subsequent items until the buffer is full. Most crucial are the model’s explanations of why some variables have different effects on the prerecency (primacy and middle) and recency portions of the serial position function. It is assumed that when a

---

2 Note that STM and LTM refer to retention over brief and long time intervals, respectively, whereas STS and LTS refer to hypothetical temporary and permanent memory systems, respectively.
A distractor task, such as mental arithmetic, is interpolated between list presentation and recall, the most recent items are no longer in the buffer at the time of recall. Indeed under standard conditions, the prerecency items are unaffected but the recency effect is eliminated by a distractor task. In contrast, presentation rate and list length affect the prerecency, but not the recency, portions of the serial position function. A fast presentation rate leads to lower levels of recall for prerecency positions because rapidly presented items remain in the buffer for a shorter time and, thus, less information is transferred about them to LTS. The level of recall for prerecency positions is lower for a given item in a long than in a short list because subjects are assumed to make a fixed number of searches of LTS, so that the probability of retrieving a particular item is lower when there are more items. The recency items are not affected by either rate of presentation or list length because they are recalled largely from the buffer rather than from LTS.

Early problems for the Atkinson & Shiffrin (1968) model were raised by Craik & Lockhart (1972), who proposed instead a levels-of-processing framework. The central postulate of that framework is that information is encoded to different levels and that the level of processing determines the subsequent rate of forgetting. A distinction is drawn between Type I processing, which maintains an item at a shallow level, and Type II processing, which promotes a deeper level of encoding. Demonstrations that maintaining information at a shallow level of processing do not necessarily lead to enhanced long-term retention were taken as evidence against the Atkinson & Shiffrin model’s assumption that transfer to LTS is a function of time in the rehearsal buffer. However, as Raaijmakers (1993) recently pointed out, the Atkinson & Shiffrin model made a distinction between the control processes of rehearsal and coding that is analogous to the distinction between Type I and Type II processing.

**PRIMARY MEMORY**

The controversy about the distinction between primary and secondary memory still exists today, as was most evident in a series of recent articles addressing the question “Short-Term Memory: Where Do We Stand?” This current debate is best exemplified by two sets of opposing quotations from articles in this series. Crowder (1993) stated: “The popularity of short-term stores grew during a time when we were busy inventing such storage receptacles. Nowadays that attitude seems archaic and, to some of us, even downright quaint” (p. 143). In contrast, Shiffrin (1993) summarized the current situation as follows: “Over the years, a metatheoretical view of short-term memory has developed. This view, closely related to the ‘modal’ model from the 1960s, is supported by an increasing base of neurophysiological data, and a wide variety of empiri-
cal findings…. The main problem with this view is the fact that it encompasses virtually everything that we are concerned with in human cognition—a successful model would almost be a general model of cognition” (p. 193).

Crowder’s (1993) empirical challenges to the concept of STM concerned both the Brown-Peterson distractor task and the recency effect in free recall. He pointed to two observations that are inconsistent with the idea that primary memory is responsible for the rapid forgetting observed across the retention interval in the distractor task. First, Keppel & Underwood (1962) found that there is no forgetting across the retention interval on the first trial of a series of trials. Second, Turvey et al (1970) showed that there is no effect of delay on forgetting when the retention interval is varied between, rather than within, subjects.

Crowder (1993) also pointed to three observations that are inconsistent with the idea that primary memory is responsible for the recency effect in free recall because of recency effects found in LTM. First, Roediger & Crowder (1976) demonstrated a recency effect in students’ recall of the names of the United States presidents. Second, Baddeley & Hitch (1977) found a recency effect for rugby team members’ recall of the teams they played against that season. Third, Bjork & Whitten (1974) discovered a recency effect in the continuous-distractor paradigm, in which a distractor task follows the presentation of each item in the list including the final item, so that primary memory could not affect recall performance. Later, Koppenaal & Glanzer (1990) changed the distractor task in the continuous-distractor paradigm after the last list item and found depressed performance, suggesting that recency does reflect a temporary rehearsal buffer. However, Neath (1993a) showed that changing the distractor task depressed performance even for prerecency items (see also Thapar & Greene 1993).

Some of these empirical challenges to the modal model can be dismissed by two general considerations. First, information can be encoded in secondary memory even with rapid stimulus presentation, so that retention can derive from secondary memory as well as primary memory in the distractor task. For example, Keppel & Underwood’s (1962) finding of no forgetting on the first trial in the distractor paradigm is easily understood in terms of the lack of proactive interference on secondary memory in the first trial. Thus, secondary memory can support first-trial retention. Second, not all bowed serial position functions have the same specific shape or the same underlying causes. For example, the shape of the serial position functions for the recall of presidents’ names has a much larger recency effect than is typically found in free-recall tasks of episodic memory, and frequency of exposure to the different presidents’ names readily accounts for this serial position function. Likewise, Raaijmakers (1993) recently pointed out that the modal model does not assume that STS is the cause of all recency effects. An advantage for recency items can
also be predicted by the modal model from the fact that retrieval from LTS is
often based on cues from the current context, and the recency items are more
closely linked to the current context than are earlier items.

*Modern Extensions of the Modal Model*

Although the original modal model continues to have its proponents, it has
been revised and extended in recent theoretical developments. We focus here
on two of these modern extensions of the modal model that maintain as
fundamental the distinction between primary and secondary memory. The
first, the Search of Associative Memory (SAM) model (Raaijmakers & Shiffrin
1981, Shiffrin & Raaijmakers 1992), is a direct descendent of the Atkin-
sion & Shiffrin (1968) model and has most extensively addressed free recall.
The second, the perturbation model (Estes 1972, Lee 1992), is primarily aimed
at performance in the Brown-Peterson distractor task using the specific para-
digm introduced by Conrad (1967).

**SEARCH OF ASSOCIATIVE MEMORY MODEL.** The SAM model elaborates the
Atkinson & Shiffrin (1968) model primarily in its description of the search and
retrieval processes from LTS that occur at the time of the memory test. At the
heart of the SAM model is the idea that events are stored in memory as images
(i.e., as separate, unitized representations) and that LTS is accessed via retrieval
cues. The strength of a given cue in terms of its link to a given memory image
is determined both by preexisting relationships and by rehearsal and coding
processes conducted in STS. For example, the link strength between an item cue
and the image of another item in the same list depends in part on the time that
the two items were rehearsed together in the STS buffer. The three most
important types of retrieval strength include self-strength (from an item as a cue
to itself as a target), associative strength (between different items), and context
strength (for linking the context cue to a target item).

SAM has been successfully fit to many aspects of free recall, paired associ-
ate recall, and interference paradigms, as well as to recognition. For recall, the
subject is assumed to generate cues at each stage of the search, starting with
the context cue and then employing other types of cues, with weights assigned
to the different cues on the basis of their salience. Because the sum of the
weights is assumed to be limited, the weights can be viewed as reflecting the
limited capacity of STS at the time of retrieval. The cues combine multiplica-
tively for recall, which allows the search process to be focused on those
memory images that are most strongly linked to all of the cues. Each retrieval
attempt involves sampling one memory image based on its strength relative to
that of all other images stored in LTS. A search termination occurs on the basis
of unsuccessful search cycles, although a rechecking process sometimes fol-
ows initial termination. After sampling an image of an item, recall of that item
depends on its recovery, which is an exponential function of the sum of the weighted strengths of the cues to the image. Most impressive is the fact that SAM can account for the part-list cuing effect in free recall, which is the finding that the likelihood of recalling a given list item decreases when other list items are given as test cues (Raaijmakers & Shiffrin 1981). SAM’s explanation of the part-list cuing effect, unlike some others, can also account for its reversal in a delayed testing situation (Raaijmakers 1993).

Recognition in the version of SAM used by Gillund & Shiffrin (1984) does not occur as a search but rather as a function of global activation (but see Mensink & Raaijmakers 1988 for a different approach). Although for recall the overall strength of a set of cues to a particular target depends on the product of the individual cue strengths, recognition is largely a direct-access process dependent on the sum of activation strengths. This difference between recall and recognition retrieval processes is central to the explanation of the list-strength effect (Ratcliff et al 1990, Shiffrin et al 1990). Strengthening some items in a list decreases the free recall of other items but has either no effect or a positive effect on the recognition of other items. The interfering list-strength effects are, therefore, attributed by SAM to retrieval processes occurring in free recall but not in recognition. Murnane & Shiffrin (1991a) take the list-strength findings as crucial evidence against memory models that postulate that structural interference occurs when storing multiple inputs in memory, such as composite storage models (e.g. Murdock 1982). Unlike SAM, these models do not assume a distinct localized memory trace or image but assume instead that a memory trace is part of a combination or superimposition of multiple traces. Ironically, however, the crucial assumption SAM makes to account for the list strength effect involves a type of composite storage; it is assumed that repetitions of the same item in the same context are stored in a single memory image (Murnane & Shiffrin 1991b). SAM must also make an additional assumption, referred to as the “differentiation hypothesis,” according to which the activation produced by an unrelated item cue is less on a stronger image than on a weaker image, presumably because the stronger image is easier to differentiate from the cue (Shiffrin et al 1990).

Although SAM is a descendent of the modal model, some of its more recent assumptions resemble those of the alternative levels-of-processing framework. For example, Shiffrin et al (1989) made the point that the units of storage used by SAM depend on coding operations, and longer and deeper operations result in higher-order units. Also, in accord with the notion of encoding specificity, Clark & Shiffrin (1987) postulated context sensitive encoding, which reflects the fact that the coding of a particular item is influenced by the other items in the same group, and test performance is best when the group of test items matches the group of study items.
Although SAM is able to account for the full range of accuracy findings in explicit, episodic memory paradigms, it has not yet been applied to fit reaction time data in recognition, or to many implicit memory and semantic memory tasks (but see Raaijmakers 1993 for some insights into how SAM could be applied in those situations). Other challenges for SAM include complete explanations of the mirror effect in recognition memory (whereby the recognition of new items as unfamiliar, mirrors the recognition of old items as familiar; see Glanzer & Adams 1990) and the learning of crossed-list associates (e.g. hot-slow, fast-cold; see Humphreys et al 1989).

PERTURBATION MODEL. The perturbation model (Estes 1972, Lee & Estes 1981) was designed to account for four principal findings from the Brown-Peterson distractor paradigm. First, there is a steep retention function reflecting rapid forgetting. Second, there are symmetrical bowed serial position functions at each retention interval in the recall of order, but not item, information. Third, there are gradually declining positional uncertainty curves, reflecting the fact that a letter substituted for another at recall usually comes from a neighboring serial position. Fourth, when to-be-remembered items are divided into separate segments, there is a recall advantage for the most recently presented segment.

According to the perturbation model, codes for immediate recall of order information are arranged in a hierarchical structure containing multiple levels. These levels include the position of the item within a segment, the segment containing the item, and the specific trial on which the segment occurred. The hierarchy of codes is repeatedly reactivated, and at each reactivation there is some probability that the relative position of neighboring items, segments, or trials will be transposed (perturbed). In the original version of the perturbation model, there was a single free parameter, theta, the probability of a perturbation at a given level of the hierarchy in primary memory. In a later version (Healy et al 1987), a second parameter, alpha, was added to the perturbation model. This parameter is the probability that a memory code will be subject to the perturbation process; hence, 1-alpha is the probability of storing position information in secondary memory. In other words, the perturbation model now has two free parameters, the first reflecting primary-memory rehearsal processes and the second reflecting secondary-memory encoding processes.

The necessity for including the secondary-memory parameter was demonstrated in a series of experiments by Cunningham et al (1984) and Healy et al (1987) who compared two conditions, both involving the presentation of two segments of items on a given trial but the recall of only one segment. Subjects were either told in advance (precue), or at the end of the distractor task (postcue), which one of the two segments was to be recalled. Performance was consistently above the chance level and better in the precue condition than in
the postcue condition, even at retention intervals up to 30 sec. It was found that the perturbation model was able to fit the data only when the secondary-memory parameter was included. Of particular note is the fact that only the secondary-memory parameter varied in the data fits; the primary-memory parameter was constant throughout these data fits (and equal to the value used by Lee & Estes 1981).

The constant value of the primary memory parameter across experiments implies that there is a fixed rate of forgetting from primary memory that can be measured using the Brown-Peterson distractor paradigm. However, Muter (1980) argued that the rate of forgetting from primary memory is underestimated by the Brown-Peterson paradigm because, in a typical experiment using the distractor task, subjects are repeatedly tested, so that they develop a high expectancy to recall after the distractor period. Muter offered an alternative paradigm in which subjects were led to expect either to recall immediately or to perform a distractor task without subsequent recall. Subjects were tested following a distractor-filled retention interval on only a few critical trials. Muter, and subsequently Sebrechts et al (1989), found a dramatic increase in forgetting rate on the critical trials such that the probability of correct recall approached floor-level performance after only 2–4 sec. However, more recently, Cunningham et al (1993) used two new versions of the distractor task to reduce the involvement of secondary memory processing. This reduction was achieved either by manipulating subject expectancy to recall or by manipulating the importance of the to-be-remembered material. With both methods, recall performance was substantially depressed at all retention intervals (including immediate recall); there was no evidence of a more rapid forgetting rate, only a change in the asymptotic performance level. Cunningham et al found that the perturbation model provided a close fit to the obtained data when they kept the primary-memory parameter at the level used in previous experiments and varied only the secondary-memory parameter such that the involvement of secondary memory processes was minimized. These results indicate that the distractor task remains a valid paradigm to study primary memory and the perturbation model provides a valuable tool for describing forgetting from primary memory.

Although the perturbation model was developed to account for STM processes, Nairne (1991) argued that the model can also be applied to situations involving LTM. In particular, Nairne showed that the model with just the single parameter, theta, for the perturbation rate could explain the results of experiments involving the reconstruction of order information from LTM. The model accounted for the symmetrical bowed serial position functions and the gradually declining positional uncertainty curves for both list selection and placement of an item in a list.
Subsequently, Nairne (1992) reported an experiment that examined the retention of position information over intervals ranging from 30 sec to 24 h and compared the data to predictions based on the perturbation model. In applying the perturbation model to this situation, which spanned both STM and LTM, Nairne again used the version of the model with only the single parameter theta reflecting perturbation rate. However, he inadvertently added another parameter, because he assumed that there was an opportunity for perturbation once every 6 sec in the first 30-sec interval but once every 24 min thereafter. In other words, the forgetting rate was very rapid over the period corresponding to immediate retention but was considerably slower over longer retention intervals. This observation suggests that a two-parameter model (one for primary memory and another for secondary memory) is necessary. In fact, even with the two separate forgetting rates in Nairne’s application of the perturbation model, the observed proportion of correct responses was significantly greater than the predicted proportion for the 24-h retention interval.

Although the two-parameter version of the perturbation model can be viewed as an extension of the modal model because it maintains as fundamental the distinction between primary and secondary memory, Estes (1991) challenged another underlying assumption of the modal model, which he terms the “trial-unit” assumption, by which each trial is considered as a discrete episode. Estes observed that item intrusions on a given trial are derived largely from the recall responses of the previous trial. These item intrusions were given in the same position as they occurred in the previous recall output, whether or not that position matched the previous stimulus input. Items from previous trials that were not recalled on those trials usually did not intrude into the recall responses of the given trial. These findings were taken as evidence against the trial-unit model and in favor of an alternative continuum model, in which there are no strict boundaries separating representations of items from successive trials. More recently, Estes (1994) provided a formal account of this phenomenon in terms of his array model of classification, with the assumption that each to-be-remembered item on a trial includes a context feature which it shares with the other items from the same trial, but which differs from those of items from previous trials, with the difference increasing with greater distance between the trials.

To examine intrusions of responses from rehearsal on a previous trial, Estes (1991) included a condition in which to-be-remembered letters were followed by distractor digits and then recalled after a delay during which subjects rehearsed the letters aloud. He found that items rehearsed out of the correct input position were less likely to be recalled than were those rehearsed in the input position. Also, when the recall and rehearsal positions were not the same, the majority of the items were recalled in their correct input position. This finding suggested that recall order was not a simple repetition of rehearsal.
Unrehearsed items were likely to be recalled in their input position or in a neighboring position, although the recall occurred with a lower probability and with less precise positional information than that for rehearsed items. To explain these observations, Estes contrasted the direct recall track, from input to recall, with an indirect track, from rehearsal to recall. The indirect track is stronger because the rehearsal context, relative to the input context, is more similar to the recall context. By postulating two different recall tracks, Estes seems to be moving towards memory models with multiple STM systems. We turn next to a discussion of models of this type.

Additional Memory Systems

Although the modal model has traditionally been described as containing a single primary memory store, the Atkinson & Shiffrin (1968) version did distinguish between the rehearsal process and other control processes in STS. This distinction foreshadowed subsequent theorizing that has broken down primary memory into separate systems.

Working Memory

Undoubtedly the most influential model that analyzes primary memory into separate systems is Baddeley’s working memory model (Baddeley 1992), which emphasizes the active processing rather than the passive storage of information that occurs in primary memory. Because of experiments using a dual-task technique in which a concurrent memory load of three or six digits had no influence on the recency part of a free-recall test for words but had small effects on performance of reasoning and prose comprehension, Baddeley & Hitch (1974) argued against a unitary limited-capacity STS. They proposed instead a multicomponent working memory model, including a limited-capacity attentional system, the central executive, supported by two slave systems, the phonological (or articulatory) loop and the visuo-spatial sketchpad (or scratchpad).

The phonological loop has received the most empirical support. It contains a phonological (i.e. speech-based) store along with an articulatory control process. The store maintains information by means of subvocal rehearsal, and without such rehearsal, information rapidly decays from the store over a period of seconds. There are at least six robust effects that are explained in terms of the phonological loop (Baddeley 1992). First is the phonological similarity effect, whereby immediate recall is lower for items that are phonologically similar than for those that are dissimilar, presumably because the similar material contains fewer phonologically distinctive features. Second is the irrelevant speech effect, whereby spoken material irrelevant to the memory task and from a different speaker disrupts immediate memory performance, presumably because all spoken material automatically enters the phonological store irrespective of its meaning (Salame & Baddeley 1989). Third is the
phonological sandwich effect, in which irrelevant spoken material is interpolated between to-be-remembered items. This effect is like that of irrelevant speech and has been explained in a similar manner, but both the list and interpolated items are by the same speaker in this case (Baddeley et al. 1991). Fourth is the word-length effect, whereby immediate serial recall of words depends on their spoken duration, presumably because shorter words are rehearsed more rapidly than longer words. In fact, it has been shown that subjects recall roughly as many words as they can say in 2 sec (Baddeley et al. 1975). Fifth is articulatory suppression, whereby immediate memory for visually presented material is depressed by requiring subjects to articulate irrelevant material during stimulus presentation, presumably because the irrelevant material blocks the articulatory control process and prevents other material from entering the phonological store. Supporting this description is the finding that articulatory suppression eliminates the phonological similarity, irrelevant speech, and word length effects (Baddeley et al. 1984). Sixth are observations involving patients with deficiencies in immediate memory who do not show phonological similarity or word-length effects (Vallar & Shallice 1990), and patients with deficiencies in speech articulation but unimpaired language processing who do show such effects (Baddeley & Wilson 1985), presumably because the former patients, but not the latter, have a defective phonological store.

Questions have been raised concerning the usefulness of the phonological loop (Baddeley 1992) because patients with phonological deficiencies appear to have normal functioning outside the laboratory and show only minor problems with sentence comprehension (Butterworth et al. 1986). However, such a patient was shown to have a marked problem learning Russian vocabulary (Baddeley et al. 1988); hence, the phonological loop does seem to play an important role in long-term phonological learning. Also, Gathercole & Baddeley (1990) found that subjects with delayed development of language had a reduced-capacity phonological loop. With normal subjects, a deficit in the phonological loop can be simulated by the use of articulatory suppression, which, similarly, has been shown to disrupt new phonological learning but not paired-associate learning (Papagno et al. 1991).

There has been less progress in understanding either the visuo-spatial sketchpad or the central executive. Secondary tasks, ranging from nonvisual spatial tracking tasks (Baddeley & Lieberman 1980) to nonspatial visual observation tasks (Logie 1986), have been used to pinpoint the nature of the processing of the sketchpad. These tasks have shown that the sketchpad involves both spatial and visual processing and is distinct from the verbal processing associated with the phonological loop. The central executive is an attentional control system that is in charge of both strategy selection and integration of information from various sources, including the two slave sys-
tems. Baddeley (1992) pointed to randomly generating letters as a useful secondary task to load the central executive, neuropsychological studies of frontal lobe damage as providing valuable insights into the operation of the central executive, and a model of attentional control as the most promising candidate for the core of the central executive.

Baddeley (1992) showed how his tripartite system can illuminate the cognitive functions affected by various diseases and those used in various tasks. For example, various secondary tasks were used to disrupt the central executive and the two slave systems in normal and Alzheimer’s disease patients (Baddeley et al. 1986, Spinnler et al. 1988). The results suggested that the Alzheimer impairment lay with the central executive rather than the slave systems. Likewise, studies using a similar methodology to analyze chess performance provided evidence that both the visuo-spatial sketchpad and the central executive play important roles in coding and memory during chess but that verbal coding is not used.

Baddeley & Hitch (1993), while retaining their working memory model, have proposed a new explanation for the recency effect in free recall in terms of a simple temporal or ordinal discrimination hypothesis, so that recency reflects registration in implicit memory by a priming process. They proposed that recall involves reactivating nodes for items in a network and that more recent items are primed and thus reactivated more easily. In support of their hypothesis of recency as priming, they assessed recency in terms of Tulving & Schacter’s (1990) five criteria for priming: They pointed to (a) intact performance in amnesia for recency items, (b) developmental dissociation between recency items and the rest of the list, with no developmental effect on recency items, (c) drug dissociation between recency items and the rest of the list, with no effect of drugs on recency items, (d) functional independence, whereby recall of earlier list items is influenced by different factors than is recall of recency items, and (e) at least some hint of stochastic independence (but see Hintzman & Hartry 1990 for issues concerning stochastic independence), whereby initial recall of recency items is independent of their subsequent recall from LTM. Baddeley & Hitch argued further that although recency reflects implicit learning, it depends on the use of an explicit retrieval strategy in which the last items in the list are output first.

**CONCEPTUAL SHORT-TERM MEMORY** Although Baddeley’s (1992) phonological loop seems to play little role in sentence understanding, Potter (1993) has proposed a distinct conceptual STM (CSTM) that is not evident in the standard paradigms for studying STM but plays an important role in everyday reading, scene perception, and sentence processing. This conceptual memory is laid down and decays very rapidly (within 1 sec). When a visual stimulus is presented, such as a sentence, the conceptual output goes to CSTM and the
phonological output to conventional STM (i.e. the phonological loop). CSTM holds considerable information, most of which is rapidly lost unless it is relevant to a conceptual structure that is consolidated into LTM.

Potter has used the technique of rapid serial visual presentation (RSVP) of pictures or words to a fixed location on a computer screen to study CSTM. In one set of her studies (Potter 1976), RSVP pictures were presented either with immediate rapid semantic detection (e.g. respond to a picture of a picnic) or a recognition test following the sequence. Subjects performed well on the detection task, but performance was near chance levels on recognition. This finding is consistent with rapidly decaying CSTM. Likewise, Potter (1993) reviewed a study involving RSVP with words, in which words could be identified in two-word lists at very rapid rates (faster than 3 words/sec), but those same rates yielded memory span performance for longer lists well below that in the standard STM task with slower presentation rates (e.g. 1 word/sec). When the RSVP lists consisted of sentences (Potter et al 1986), even longer lists of words were recalled nearly perfectly even at high rates, suggesting that they were parsed and understood immediately. Thus, words at fast rates enter Potter’s CSTM but not conventional STM.

Lombardi & Potter (1992) and Potter & Lombardi (1990) proposed that immediate recall of RSVP sentences involves regeneration from a conceptual representation formed from the activated part of LTM. The verbatim nature of RSVP sentence recall reflects the fact that the lexical items in the sentence were activated more than other lexical items. In support of this hypothesis, subjects’ sentence recall included intrusions of closely synonymous words from an interpolated secondary word-matching distractor task, even when the sentences were presented extremely rapidly. In another study involving RSVP sentence recall (Potter et al 1993), ambiguous misspelled words (e.g. *duck* which could be a misspelling of *duck* or *deck*) were replaced by subjects in their recall protocols with the correctly spelled words appropriate to the sentence context. This finding suggests that the sentence context was processed conceptually despite the rapid rate of sentence presentation.

Potter (1993) related CSTM to conceptual priming but noted that concept reactivation cannot account for the formation of episodic-specific links among word representations, as occurs with the RSVP presentation of a sentence. CSTM provides the basis for encoding information in LTM according to Potter’s account and can be viewed either as a separate episodic representation or as an activated part of LTM. In the light of this latter view, Potter’s approach can be seen as similar to an approach taken recently by Ericsson & Kintsch (1995).

**LONG-TERM WORKING MEMORY** Ericsson & Kintsch (1995) raised the question as to whether the standard definition of working memory as limited, temporary
storage can be applied to tasks such as reading text and skilled performance by experts, especially because these activities can be interrupted and later resumed without major detriments in performance. To resolve this question, Ericsson & Kintsch proposed a mechanism based on skilled use of storage in long-term working memory (LT-WM) in addition to the temporary storage of information in short-term working memory (ST-WM). This work is an extension of Chase & Ericsson’s (1982) skilled memory theory, which accounts for a tenfold increase in digit span by individuals given extensive training. This training overcame the proactive interference caused by previous storage of similar information in memory. Chase & Ericsson proposed that such individuals draw on their acquired knowledge and on systems of retrieval cues, or retrieval structures.

Ericsson & Kintsch (1995) took issue with the assumption that it is slow both to retrieve information from LTM and to store new information in LTM. Although they acknowledged that it does indeed take up to 1 sec to retrieve—and 10 sec to store—unfamiliar information in LTM, it takes much less time for an expert to retrieve or store relatively familiar information in LTM. Likewise, although the estimate of STM capacity as about four chunks of information (Broadbent 1975) applies well to a wide range of simple cognitive activities, for more complex tasks such a small working-memory capacity is insufficient (Anderson 1983, Newell 1990).

To accommodate the increased memory performance by chess experts, Chase & Simon (1973) argued that the experts relied on larger chunks in STM. However, Ericsson & Kintsch (1995) pointed to findings that chess experts can recall up to nine different patterns of chess pieces (Cooke et al 1993), and that interpolated tasks do not depress performance by such experts, as suggesting storage in LTM rather than STM. Similarly, the performance of memory-span experts seems to reflect storage in LTM because of small decrements in performance with STM interference before list recall, accurate recall of all lists at the end of a session, and improvement specific to the type of material practiced.

Ericsson & Kintsch’s (1995) account is that subjects use LT-WM for a skilled activity when they have a large body of relevant knowledge and can anticipate future memory demands. The use of LT-WM relies on a stable retrieval structure, which is all that needs to be available in ST-WM, along with a cue indicating the relevant type of information required. In contrast, Schneider & Detweiler (1988) argued that LTM cannot be used as working memory because of the build-up of retroactive interference, and Baddeley (1990) used the finding that expert mental abacus calculators can recall no more than one sequence of digits (Hatano & Osawa 1983) as evidence against the use of storage in LTM for this task. Ericsson & Kintsch addressed these issues by pointing out that temporal distinctiveness and elaborative encoding
can overcome the build-up of retroactive interference and that temporal distinctiveness is poor in the task of mental abacus calculation.

Ericsson & Kintsch (1995) used Kintsch’s (1988) construction-integration model to explain how the constructed representation of a previously read text is kept accessible in LT-WM so that new information can be encoded and integrated with previous information. Glanzer et al (1984) interrupted text reading with an unrelated task and then allowed reading to resume. The effect was to increase reading time for the first sentence after reading resumed but not to influence speed or accuracy of answers to comprehension questions. Ericsson & Kintsch took these findings as evidence that the disruptions led to a loss of retrieval cues in ST-WM; comprehension was not disrupted because the information was stored in LT-WM. They further proposed that superior text comprehension is due to superior skill encoding information in LT-WM and making it accessible via cues in ST-WM. Contrary to the assumption by Just & Carpenter (1992) that good readers have more room in active memory than poor readers (see Engle et al 1992 for a similar view), Ericsson & Kintsch proposed that good readers produce a more extensive retrieval structure for LT-WM so that their effective working memory is larger without an increase in the size of active memory.

SHORT-TERM ACTIVATION AND ATTENTION  The relationship between STM and LTM has also been explored by Cowan (1993, 1994), who pointed out that there have been two definitions for STM, one (deriving from James 1890) that is the current focus of attention, and the other that is currently activated in LTM. These two definitions have sometimes been used interchangeably, but they cannot be equivalent given evidence for activation outside of awareness, including priming (i.e. activation) from a previously attended item. Cowan represents STM as a nested subset of LTM; the currently activated features are a subset of LTM, and the focus of attention is a subset of the activated memory.

Cowan et al (1990) supported the hypothesis that STM includes both attended and unattended information by examining immediate retention for attended and unattended stimuli. Subjects silently (or by whispering) read a novel and simultaneously heard speech syllables presented through headphones. They were given comprehension tests and asked to write a sentence summarizing recently read material. On crucial test trials they were asked to recognize the speech syllables heard, after a variable delay filled with the reading. A rapid memory decay was found for the speech syllables. Performance was improved on shorter delays if subjects shifted their attention away from reading, as evidenced by their not whispering during a short interval on either side of the target syllable. If subjects consistently divided their attention between the two channels by responding to particular syllables, then little forgetting of the target syllables was evident even at the longest (10 sec) delay.
Thus, there was memory for unattended syllables, and subtle attentional shifts improved memory.

Cowan et al (1992) examined STM loss by studying memory decay during subjects’ overt verbal recall of a list. They used words identical in number of phonemes and syllables but different in pronunciation length. They found that longer words at the beginning of the list depressed forward recall, whereas longer words at the end of the list depressed backward recall. This finding suggests that while subjects recall words from the list the other words decay from activation. Cowan (1993) proposed that when a list is recalled, memory decays during word output but is reactivated during the pauses between words [see Schweickert (1993) for a candidate model of this process]. Individual differences in immediate memory span are then attributed not to the speed of pronouncing words but to the efficiency of covert processing between words. This hypothesis is supported by Cowan’s (1992) finding of (a) a correlation for four-year old children between span and duration of recall but not between span and speech rate and (b) faster spoken recall of lists less than span length, with the reduction in speed located entirely within the pauses between words.

Central to Cowan’s (1993) framework is the distinction between STM and LTM. Two findings taken as support of a separate STM are the recency effect and the word length effect in immediate recall. The recency effect in the continuous-distractor paradigm, which relies on LTM, has been taken as weakening the support for a separate STM (Greene 1986). However, Cowan et al (1994) provided evidence that the word length effect was not the same in immediate recall and in the continuous-distractor paradigm. They manipulated the syllable length of the words on a list and asked subjects for backward recall. For the immediate recall procedure, increased word length of the items in the second half of the list depressed recall, but the reverse was found (i.e. better performance with longer words) for the continuous-distractor procedure. Further, in support of Cowan’s (1993) proposal that STM is the activated portion of LTM, Hulme et al (1991) found a linear relation between immediate memory span and speech rate, with a higher intercept for words than for nonwords, so that lexical familiarity in LTM made a difference in STM.

Unitary Memory System

Although the evidence presented for additional memory systems is quite compelling, Roediger (1993) persuasively discussed the interpretive problems created by a proliferation of memory stores. In contrast, some theories postulate a unitary memory system. We will concentrate on one such theory proposed by Crowder & Neath (1991) because it was specifically formulated as an alternative to the modal model. Crowder & Neath argued that the Brown-Peterson distractor paradigm, rather than isolating STM, serves to temporally magnify the retention of a particular item, as if it were under a microscope. They used
this microscope analogy to suggest that the same principles apply to STM as to LTM. Following Murdock (1960), they focused on the principle of distinctiveness, which they used to explain serial position functions in both STM and LTM. Murdock measured distinctiveness of a serial position by computing the difference between its ordinal number and that of all the other positions in the list. In addition, he accounted for the asymmetry of the serial position function by transforming the ordinal numbers into log values (see Johnson 1991 for a recent discussion and elaboration of these ideas). This theory necessarily assumes a constant shape of serial position functions across learning conditions (e.g. variations in presentation rate and item familiarity), and indeed the serial position function for serial learning is constant when normalized (i.e. plotted as the proportion of the total number of correct responses made at each position in the list). Neath (1993b) extended Murdock’s formulation by using the actual durations of the interstimulus and retention intervals when calculating the temporal distinctiveness of an item.

Following Glenberg (1987), Neath & Crowder (1990) employed the continuous-distractor paradigm to study temporal distinctiveness. They compared a control condition with equal temporal spacing (interpolated arithmetic) between successive list items to an increasing condition, in which the spacing increased as the list progressed, and to a decreasing condition, in which the spacing decreased. They showed that for items presented visually the recency effect was largest in the increasing condition (in which the distinctiveness of the recency items was greatest) and smallest in the decreasing condition (in which the distinctiveness of the primacy items was greatest). They also conducted a parallel experiment using the Brown-Peterson distractor paradigm in which they interpolated digits between the to-be-remembered letters with a constant, increasing, or decreasing number of digits. The increasing condition showed the most recency and the decreasing condition the least. Thus, the results in the two situations were parallel despite the fact that the first relied on LTM and the second on STM.

Also consistent with this view is a study by Wright et al (1985) with data from monkeys, pigeons, and humans viewing four slides followed by a retention interval and then a recognition probe. At immediate testing, there was a pure recency effect with no primacy effect, whereas with delayed testing there was greater primacy with increasing delay until there was pure primacy and no recency. Crowder & Neath (1991) explained these findings by proposing that the retrieval orientation was initially from the recency end of the list but after a delay shifted to the primacy end. Neath (1993b) replicated these findings in studies with human subjects recognizing pictures of snowflakes that could not be easily verbalized. Whereas the duration of the retention interval was varied in this study, in a subsequent study, Neath & Knoedler (1994) varied the duration of the interstimulus interval and also found results consistent with the
distinctiveness model. Further, Neath & Knoedler applied the same model to experiments by Gernsbacher & Hargreaves (1988) and their own follow-up experiment on sentence processing. These experiments showed a response-time advantage for verifying the second-mentioned of two participants after a short retention interval (analogous to a recency effect), but an advantage for verifying the first-mentioned participant after a longer interval (analogous to a primacy effect). Thus, the distinctiveness model can account for a wide range of memory phenomena, including STM and LTM paradigms, experiments with animals and humans, and tasks ranging from picture recognition to sentence processing. Nevertheless, it is not clear to us how this model would account for the full range of empirical phenomena we have reviewed in support of multistore models.

Summary

The modal model has been successfully elaborated both in terms of secondary memory search and retrieval processes and in terms of primary memory forgetting processes. Nevertheless, there is strong empirical evidence that distinctions are necessary beyond the modal model’s dichotomy between primary and secondary memory. In particular, there is evidence for separate phonological and visuo-spatial processing and for a distinction between indirect and direct recall tracks or, similarly, between short-term activation and attention. Evidence also points to an additional very short-term conceptual memory and a more active involvement of LTM in skilled performance. The concept of distinctiveness accommodates a wide range of findings previously attributed to primary memory, but it seems doubtful that it can provide a comprehensive account.

SENSORY MEMORY

We now consider the front end of the modal model: sensory memory. The modal model includes separate memories for each sense, including visual, auditory, and haptic. Because visual sensory (iconic) memory has been shown to be useful for only a fraction of a second [and perhaps not useful at all outside artificial laboratory tasks (see Haber 1983)], and because haptic and other senses play little role in verbal memory tasks, we concentrate here only on auditory sensory (echoic) memory.

Precategorical Acoustic Store

The strongest evidence for a distinct auditory sensory memory came from three interrelated observations in immediate serial recall of auditorily presented words (usually digits): (a) the recency effect (the advantage for the last item of the list), (b) the stimulus suffix effect (the elimination of the recency
effect with the presentation of a redundant word presented after the list as a recall cue), and (c) the modality effect (the elimination of the recency effect with visual presentation of the list). These findings prompted Crowder & Morton (1969) to postulate an auditory sensory memory, which they termed precategorical acoustic store (PAS); their description of PAS was subsequently extended by Crowder (1978) to include effects of auditory masking. The store was labeled precategorical because it was found to be sensitive to physical (e.g. the voice of the speaker) but not semantic (e.g. the meaning of the words) characteristics of the items. The store was labeled acoustic because it seemed to depend on auditory presentation of the items. However, both of these characteristics of PAS have been disputed in recent years.

The precategorical nature of PAS was disputed by findings indicating that the suffix effect depended on the subjects’ interpretation of the suffix as either speech or nonspeech. For example, Neath et al (1993) used as a suffix the word “baa.” Half of the subjects were told that it was an actual sheep sound while the other half of the subjects were told that it was produced by a person. The suffix effect was observed only for the subjects who thought that the suffixes were spoken by a human. The acoustic nature of PAS was disputed because of studies showing that silently mouthed or lipread stimuli produced suffix effects even when the lists of to-be-remembered items were heard (Greene & Crowder 1984).

Both of these problems for PAS can be resolved by viewing it as a speech memory rather than an acoustic memory because it is well known that visual articulatory information, such as lip movements, affects speech perception (McGurk & MacDonald 1976) and that the interpretation by listeners of a specific sound as speech or nonspeech affects the way it is perceived (Liberman 1982). Work in speech perception has also revealed important differences in the perception of vowels and stop consonants (Healy & Repp 1982). This revised PAS account (Crowder 1983) can, thus, easily accommodate the fact that vowels but not stop consonants show suffix and modality effects for auditory (Crowder 1971) and lipread (de Gelder & Vroomen 1994) stimuli (but see Turner et al 1987 for different findings with mouthed stimuli).

An important challenge to this revised account of PAS is that modality effects have also been reported in paradigms involving LTM, such as the continuous-distraction paradigm (Gardiner & Gregg 1979). However, Greene (1992) pointed out that these long-term modality effects (unlike the parallel short-term effects) are not found for serial recall of digits. Hence, it may not be necessary to use the same explanation for both short-term and long-term modality effects. Nevertheless, alternative theoretical frameworks have been proposed to account for the full range of modality effects. One such account is Glenberg’s (1987) temporal distinctiveness theory, which is based on the assumption that temporal discrimination is more accurate for auditory than for
visual presentation of items. In support of this distinction between auditory and visual information, Glenberg et al. (1989) found an advantage for auditory, relative to visual, presentation in the reproduction of temporal rhythms. On the other hand, Schab & Crowder (1989) disputed this distinction because they found little auditory advantage in the estimation of temporal durations.

**Processing Streams**

Penney (1989) also rejected the PAS account of the short-term modality effect and proposed an alternative account in terms of separate processing streams for auditory and visual items. According to her account, auditorily presented items automatically elicit both A (acoustic) and P (phonological) codes. The A code is sensory-based and produced only for stimuli that are heard. The P code is internally generated and is analogous to Baddeley’s (1992) phonological loop, including information about words, phonemes, and articulation. Visually presented items are normally represented in the P code, but not automatically so. In addition, they elicit a visually-based code analogous to Baddeley’s (1992) visuo-spatial sketchpad. Penney’s support for the separate streams involves five different lines of evidence: (a) less interference with two concurrent verbal tasks when two modalities, rather than a single modality, are employed, as in attentional studies; (b) improved memory when list items are presented in two different modalities; (c) modality selective interference effects, including suffix effects and effects of different distractor tasks; (d) the advantage for recall organized by modality, as opposed to time of presentation; and (e) STM deficits that are modality specific.

Penney (1989) raised two primary criticisms against the PAS model. The first concerns its underestimation of both the capacity and duration of auditory sensory memory. Penney argued that the A code represents at least five items and lasts as long as a minute. Second, and most importantly, she pointed to demonstrations of long-term modality effects, such as that found by Gathercole & Conway (1988) extending throughout a list of 30 items. Because long-term and short-term modality effects react differently to some variables, she proposed that these two effects reflect different mechanisms but that both reflect properties of the A code. Three properties of the A code emphasized by Penney are (a) its large capacity and the fact that it does not decay but is subject to interference, (b) its specialization for coding sequential associations between items, and (c) the automatic nature of its generation and maintenance. However, with Penney’s proposed large capacity for the A code, it appears difficult to provide an explanation for the typical modality and suffix effects, which are limited to the final position in immediate serial recall.
Feature Model

Nairne’s (1988, 1990) feature model is an alternative framework in which to interpret recency effects in serial recall. Nairne postulated two types of memory trace features—modality independent, which coincide with the inner voice, and modality dependent, which reflect perceptual aspects of the stimuli. The modality-dependent features are analogous to Penney’s (1989) A code and visuo-spatial sketchpad code, whereas the modality-independent features are analogous to her P code. Nairne postulated similar modality-dependent features from speech, lipread, and mouthed stimuli, which are all output of a language-analysis system. Thus, the modality-dependent features are perceptual, not sensory. In contrast, the modality-independent features correspond to STM or articulatory coding.

According to Nairne (1988), particular trace features are used at recall if they are discriminable (distinctive) and salient (useful or relevant). Recall is based on a reconstructive process in which the features of the memory trace are compared with features of candidate items. There is no decay as a source of forgetting in this account (as there is in the PAS account). Rather, forgetting results purely from overwriting, which occurs feature by feature, such that a subsequent feature can overwrite an earlier feature only if it is the same type (modality-dependent or modality-independent). Overwriting also occurs only for items in the same perceptual group (Frankish 1989). Recency results because the final list item does not suffer from overwriting of modality-dependent features unless there is a suffix. The modality-independent features of the final list item are overwritten by rehearsal and other inner voice activities that occur at the end of the list even when there is no suffix.

An important aspect of Nairne’s (1988) model is the fact that he is able to account for the presence of visual recency effects, which occur for deaf subjects with American Sign Language (Shand & Klima 1981) and for hearing subjects with abstract (Broadbent & Broadbent 1981) and unusual (Campbell et al 1983) visual stimuli. According to Nairne, auditory features are not stronger, more distinctive, or more durable than visual features. Visual stimuli usually fail to exhibit recency because of overwriting by visual events after the end of the list and because subjects generally use auditory, rather than visual, features as discriminative cues in recall. Nairne’s account is, thus, similar to Shand & Klima’s (1981) primary linguistic code account, according to which recency effects occur when the stimuli are presented in the modality consistent with the dominant format used for STM coding. However, by Nairne’s account, subjects may be led by some task demands to attend to linguistically irrelevant visual features. If the salience of visual features is enhanced in a particular task, then visual recency effects should occur according to Nairne’s model, but not according to the PAS model.
Nairne (1988) accounts for long-term modality effects in a manner similar to that of Glenberg’s (1987) temporal distinctiveness theory. Time of occurrence is associated with a broader temporal region for visual traces because they are largely composed of modality-independent features, as are the neighboring traces for inner-voice activity.

Summary
Just as attempts were made to use distinctiveness as an explanation for previous findings attributable to primary memory, it has been used as an explanation for previous findings attributable to auditory sensory memory, especially because of parallel effects found in LTM. Nevertheless, there remains powerful evidence for a separate type of processing associated with speech perception; that is, there is support for a distinction between acoustic and phonological codes or, similarly, between modality-dependent and independent features.

SECONDARY MEMORY
According to some versions of the modal model, including the simple version of the Atkinson & Shiffrin (1968) model that was applied to free recall, storage in secondary memory is permanent, with forgetting attributable solely to retrieval, rather than storage, failures. This proposed durability of storage in secondary memory is in sharp contrast to the rapid loss of information from primary memory. Hence, we concentrate our discussion of secondary memory on the issue of its permanency and durability. (For more inclusive summaries of the literature on long-term retention, see Healy & Sinclair 1995, Schmidt & Bjork 1992.)

Permastore
Although it is difficult to study retention over long delay intervals in the laboratory, an ingenious naturalistic cross-sectional method was developed by Bahrick (1984) to study retention over periods up to 50 years. This method employs a large number of subjects who acquired the same knowledge at different times in the past. These subjects must estimate the degree of their original acquisition of the material and the extent to which they rehearsed the material after acquisition. Retention tests are performed and a retention function is calculated on the basis of the subjects’ date of original acquisition and then corrected, through multiple regression, for such factors as degree of original acquisition and extent of rehearsal. Bahrick used this method to study the retention of Spanish learned in the classroom by more than 700 subjects whose last exposure to a Spanish course was from 0 to approximately 50 years prior to the retention test. Rehearsal was very low and not a good predictor of performance on the retention test, which was instead predicted to a large extent
by original acquisition level and training in other Romance languages. The retention function indicated an exponential decline in performance across the first 6 years, after which there was a stable asymptote for about 30 years, followed by a final decline probably attributable to aging. Because so much knowledge was maintained across the long delays with little intervening rehearsal, Bahrick introduced the concept of permanent memory, or “permas-tore.”

Although Bahrick’s (1984) findings have been disputed (Hintzman 1993), more recent work has provided converging evidence for the notion of permas-tore. Using the same techniques, Conway et al (1991) assessed retention over a 12-year period of material learned in a cognitive psychology course. After a rapid decline in performance across the first four years, there was a stable asymptote for the remaining eight years. Further, in a cross-sectional study of bilingual Hispanic immigrants, Bahrick et al (1994) found essentially no loss of Spanish knowledge over a 50-year period of residence in the United States. It is important to note that the subjects in this study were at least 10 years old when they entered the United States and, thus, had a solid foundation in Spanish before immigrating and that most of them continued to speak about as much Spanish as English after entering the United States. Using a different method, in which subjects were given recognition tests for former single-season television programs, Squire (1989) supported the finding that considerable knowledge may be maintained in memory for a lifetime even in the absence of rehearsal. However, he also found a gradual and continuous loss in performance, and no evidence of a stable asymptote. This finding is consistent with demonstrations that forgetting functions follow a power law (Anderson & Schooler 1991, Wixted & Ebbesen 1991). Squire attributed the discrepancy between his findings and those of Bahrick to the higher degree of learning and the greater degree of internal organization for the learned material in Bahrick’s studies.

**Procedural Reinstatement**

Remarkable durability of memory has also been found in laboratory studies of skill acquisition and retention (Healy & Bourne 1995). For example, in an experiment in which subjects were given extensive training in single-digit multiplication, Fendrich et al (1993) found considerable improvements in the speed with which subjects provided the answers to the problems across 12 sessions of training and no deterioration in performance speed across delay intervals up to 14 months after training. This high degree of skill retention was coupled, however, with a striking specificity in the skill acquired. Rickard et al (1994) found that extensive training on specific multiplication and division problems (e.g. \( _\times 7 \)) did not transfer to parallel problems with the complementary operation (e.g. \( 28 = _\times 7 \)). To account for the observed skill
durability and specificity, Healy et al (1992a) proposed a procedural reinstatement principle, according to which superior long-term retention results when the procedures (the motoric, perceptual, and cognitive operations) used at the time of acquisition are reinstated (duplicated) at the time of the retention test. This procedural reinstatement principle draws on the distinction between procedural and declarative knowledge proposed by Anderson (1983) as well as the notion proposed by Kolers & Roediger (1984) that memory representations cannot be divorced from the procedures used to acquire them.

The procedural reinstatement principle can be divided into two hypotheses, one concerning the use of procedures during study of material and the second concerning the use of the same procedures during the retention test. Support for each of these hypotheses was provided in studies of the generation effect (better memory for material that is generated than for material that is simply read by subjects) in episodic memory for the list of answers to simple arithmetic problems shown during an experimental session. The relevant procedures in this memory task are the arithmetic operations linking the problems to the answers. Crutcher & Healy (1989) showed that subjects’ memory for the list of answers to multiplication problems was enhanced when subjects were required during study to perform the multiplication operations themselves as opposed to reading the answer with the problem. Generating or reading the answer per se was not crucial, because reading and verifying an answer did enhance performance, whereas using a calculator to generate the answer did not. McNamara & Healy (1995b) showed that enhanced memory for the list of answers to arithmetic problems only occurred when subjects were able at the retention test to recall the problem operands encountered during study and use them as retrieval cues for the list of answers. That is, a retention advantage was only found when subjects were able to reinstate at test the arithmetic procedures performed at study. The procedural reinstatement account of the generation effect is similar to other accounts (e.g. McDaniel et al 1990). However, the procedural reinstatement account has the advantage of explaining a broader range of findings concerning the generation effect (e.g. McNamara & Healy 1995a, Roediger & McDermott 1993) as well as findings in other domains, including long-term retention of nonverbal as well as verbal material.

Summary
The modal model generally assumes that secondary memory is permanent. Evidence for durability comes from naturalistic cross-sectional studies of knowledge acquisition and from laboratory studies of skill acquisition. Procedural reinstatement has been used to explain both durability and specificity of training.
CONCLUSION

Does the modal model still work? Some of the alternative frameworks, including distinctiveness and procedural reinstatement, can explain a wider scope of memory findings, including those spanning both short and long retention intervals and both verbal and nonverbal material. However, it seems clear from the studies reviewed here that these alternative frameworks cannot account for the full set of findings used to support the modal model. It is also clear, though, that the modal model needs to be elaborated along the lines reviewed here. In any event, this review has convinced us that the modal model is still useful as a means to frame the current literature on verbal learning and memory.

ACKNOWLEDGMENTS

Preparation of this chapter was supported in part by Army Research Institute Contract MDA903-93-K-0010 to the University of Colorado (Alice Healy, Principal Investigator) and a James S McDonnell Foundation Postdoctoral Fellowship Award to Danielle McNamara. We are indebted to James Parker for his invaluable help preparing this chapter and Lyle Bourne, Bob Crowder, Bill Estes, Don Foss, Bill Marmie, Bill Oliver, and Larry Pinneo for their helpful comments on earlier versions of this chapter.

Any Annual Review chapter, as well as any article cited in an Annual Review chapter, may be purchased from the Annual Reviews Preprints and Reprints service.
1-800-347-8007; 415-259-5017; email: arpr@class.org

Literature Cited

DOES THE MODAL MODEL STILL WORK? 169

When long-term learning depends on short-term storage. J. Mem. Lang. 27:586–95
Brown J. 1958. Some tests of the decay theory of immediate memory. Q. J. Exp. Psychol. 10:12–21
Conrad R. 1967. Interference or decay over short retention intervals. J. Verbal Learn. Verbal Behav. 6:49–54
Ericsson KA, Kintsch W. 1995. Long-term


Johnson GJ. 1991. A distinctiveness model of