6
HUMAN ASSOCIATIVE LEARNING

During the popular heydays of the major learning theories, the hot controversies between Hull and Guthrie and Tolman were fought out in *animal* learning laboratories. Most of the controversies surrounded theoretical interpretations of how rewards influenced learning and performance, how to conceptualize what is learned, and what were the critical ingredients for learning. These debates were carried into the animal learning laboratory because of the widespread belief that animal learning would be simpler to understand than human learning, that fundamental features of most learning phenomena could be studied with the lowly rat learning to navigate through mazes. Research proceeded on the premise that learning mechanisms were universal throughout the animal kingdom, and that simple conditioning principles discovered with lower animals would apply with only slight modifications to humans. There was enough suggestive evidence at hand to make this thesis plausible: for instance, classical GSR and eyeblink conditioning carried out with humans showed that variables of the conditioning situation such as the length of the CS-US interval had effects similar to those found with animal subjects.

THE HUMAN LEARNING TRADITION

Developing alongside this theory-dominated animal conditioning research was a body of research concerning associative learning in adult humans, specifically studies concerned with adults' acquisition of simple S-R associations and chains of associations. These were called studies of *verbal learning, sensory-motor learning,* or *skill learning.* We will review here the tradition of verbal learning, since it had a definite founder, a clear beginning, and definite paradigms, all of which have been lacking in the motor-skills research area. For reviews of the motor-skills literature, see Welford (1976) or the volume edited by Bilodeau (1966).

The verbal learning tradition began with Herman Ebbinghaus, dating from the publication of his treatise *Über der Gedächtnis* (*Concerning memory*) in 1885. Ebbinghaus began his work in the tradition of empiricism and associationism. He showed that associative learning processes, which had been topics of much speculation among philosophers, could be brought into the laboratory and measured. Considering the historical context, his achievements were remarkable: (1) He relied on objective instead of introspective reports of memory, using the relearning method and the savings score to infer retention where conscious recall failed; (2) he invented calibrated units (nonsense syllables), which supplied a limitless number of new learning materials for experimentation; (3) he challenged the established laws of association, particularly those of temporal succession, by introducing a quantitative study of remote associations; (4) he used statistical methods to summarize his findings and discuss the significance and relative magnitude of effects of several learning variables.

Ebbinghaus created a new experimental situation—namely, learning lists of nonsense syllables—in which a multitude of variables could be defined and their influence on "remembering" behaviors observed. The phenomena that Ebbinghaus discovered are still dealt with today, and his associative theories were extremely long lived. Verbal learning research has been long dominated by Ebbinghaus's work. Subsequent research has teased out the variables of verbal learning situations, measured them, and determined their laws. During the course of this research, the learning situation and paradigms underwent constant small modifications; in a sense, the task of empirical analysis was continually beginning anew. The three major verbal learning paradigms that have been studied most intensively are:

1. Serial learning. The subject learns to recite a list of items (syllables, words, digits) in a specified serial order. The recitation may be unaided or prompted serially; in the latter case, after each item-recall attempt the next item is presented as a cue for the following item in the series.

2. Free-recall learning. The subject tries to recite a list of items in *any order* she chooses, at any pace. The list of items may be presented only once or repeated several times, either in the same order or a randomly varying order for the items.

3. Paired-associate learning. The subject learns a list of discrete associations (pairs of syllables or words), denoted generically as *A-B*. The pairs are typically learned under instructions that *A* is to serve as the cue (prompt, stimulus) for recall of *B* (as the response). The responses may be either known and available (such as buttons on a keyboard) or items which themselves must be learned as units (such as foreign-language words).

These paradigms are partly defined by the units to be memorized (lists of single items versus pairs), and by the required criterion performance (e.g., ordered vs. unordered recall). The memories established by either procedure can be tested by recall, recognition, or reconstruction. Each paradigm can be used with large or small amounts of material, with short or long retention intervals. In addition, of course, the person may be learning several different lists in succession, so we can examine the influence of learning one list upon the learning or retention of another. As variables have been isolated and studied, a huge backlog of empirical information has accumulated about how humans learn in these situations. Many hypotheses have been proposed to integrate and account for the evidence on some particular question.

THEORY IN VERBAL LEARNING

During the years 1900 to 1930, studies of verbal learning were carried out largely by a group of psychologists calling themselves *functionalists* (see Chapter 9 in Hilgard & Bower, 1975). Functionalism was a loose confederation of methodological ideas, but the central goal was to perform a detailed experimental analysis of important psychological skills or tasks. Thus, the guiding idea was to dissect any given task, such as serial verbal learning, into a number of components or constituent skills, and to analyze these experimentally. The concern with empirical description led the researchers to avoid the global theories and their associated controversies that raged in the animal conditioning laboratories; instead, verbal learning researchers thought truth would be revealed by patient empirical analysis of specific learning tasks (see McGeogh & Irion, 1952).

The background theory of learning underlying verbal learning research was general associationism, supplemented in later years by some concepts from Hull's theory. Ebbinghaus started his work in the tradition of associationism. The basic idea was that remembering could be reconstructed in terms of connections (associations) among ideas, these connections were recorded into the mind (memory) by the contiguous occurrence of the two ideas in consciousness. This mental contiguity of to-be-associated ideas was allegedly caused either by the objective contiguity (in time or space) of the external events that arouse their corresponding ideas in the mind, or by the person thinking of (retrieving from memory) a second idea while considering a first one. Retrieval during this second mode of contiguity was characterized by there being a similarity or relatedness between the prompting idea and the mate it retrieves from memory. By whatever means it occurred, the contiguous experiencing of ideas *A* and *B* together in consciousness was presumed to establish an associative connection or bond between their internal representations. This association, or "line," from node *A* to node *B* could vary in strength, affecting the likelihood and speed with which *B* comes to mind when *A* is entered into consciousness. A given element may be associated with a number of different elements, denoted *A-B, A-C, A-D*, and so on. These will be ordered in strength at any one time. The stronger is the *A-B* association, the more likely it is to win out in competition with alternative responses. Learning consists in gradual strengthening of the association to the correct response, so that it will occur more surely than error responses to the stimulus. Before discussing how this theory was elaborated to deal with phenomena of the verbal learning laboratory, let us first note how this tradition dealt with the question of reinforcement in human learning, since that played such a central role in other learning theories.

Reinforcement in Human Learning

The typical verbal learning experiment is poorly designed to study motivation, reward, and the influence of these factors upon speed of learning or performance. The typical subjects are college students who are already sophisticated at motivating and rewarding themselves for learning practically any material set before them. The typical motivation for subjects to engage in the verbal learning task is provided by the experimenter's instructions and the subjects' desire to be cooperative and to work to earn their pay (or course credit points). The experimenter's instructions usually define the task, orient the subjects to it, initiate rehearsal of the material with intent to learn, define the criteria for adequate or correct performance, and sustain the subjects' performance throughout training. This has been called the *intentional learning set*, and it presumably exists as a pool of learning strategies that students have acquired through formal education by the time they reach college.

Much research has shown that the intention to learn has its main effect in initiating certain rehearsal or elaborative activities with respect to the to-be-learned material (see Postman, 1964). But if these same activities can be evoked by other orienting tasks, without the intention to learn, the subject will learn about the same amount incidentally. For instance, if subjects are asked to construct meaningful sentences for word pairs (for example, "pen-record" might be converted into a sentence like "My *pen* scratched a *record*"), they will thereby associate the pairs of words as indicated by their later ability to recall the second word when prompted with the first. This "incidental" learning is in no way improved if in addition sentence-generating subjects are also told to remember the material for a later recall test. What is important is what the subjects do with the material when they are exposed to it, not their intention to learn or the reason given to them for carrying out these elaborative activities.

The researchers in verbal learning have always adopted an empirical law of effect. They realized that the experimental situation is arranged so that the reinforcement that promotes learning is information about the correct response. Thus, a response followed by the experimenter's announcement of "Right" will increase in its probability to its stimulus; and informing the subject of the correct response has a similar effect. Thorndike assumed that the subject's response would be rewarded (cause satisfaction and strengthening) when it matched the response designated as correct, and would be nonrewarded or punished otherwise. The opposing viewpoint to Thorndike's law of effect is that information about the correct response is sufficient to promote learning, and that satisfiers after the correct response are irrelevant to learning. This view and evidence supporting it clearly has more to recommend it than does Thorndike's original view of reward in human learning. Of course, for the functionalist interested in dissecting the learning tasks, the viewpoint adopted on this reward issue was relatively immaterial. Probably because they were relatively unconcerned with issues of motivation and reinforcement, the functionalists' analysis of verbal learning tasks proceeded apace without becoming bogged down and entangled in the grand debates and controversies between the global theories (like the Hull vs. Tolman debates) that raged from the 1930s to 1950s.

Let us turn now to elaborations of the basic associative theory developed by verbal learning theorists to explain basic phenomena revealed in studies of serial learning, of paired associate learning, of transfer of training, and of forgetting.

SERIAL LEARNING HYPOTHESES

Ebbinghaus investigated only serial learning and established many functional relationships. Included were the effects of list length on time to learn a list, the effects of different lengths of study time or number of study trials upon subsequent amount retained, and the effects of the duration of the retention interval on amount retained. In an associative analysis, a serial list, denoted *A-B-C-D-E* . . ., would be represented in memory by a chain of direct associations, *A* to *B*, *B* to *C*, and so on, so that seeing or thinking of item *A* arouses the *A-B* association and produces response *B*, which then arouses the *B-C* association and produces response *C*, and so forth. Ebbinghaus, however, observed that the learners themselves made frequent errors by recalling an item at an earlier position than was correct. Thus, after recall of *A* and *B*, the subject might next misrecall *D* or *E* instead of *C*. These anticipatory errors followed a distance gradient, with near errors being more likely.

The Doctrine of Remote Associations

These distant anticipatory errors led Ebbinghaus to postulate that the associative structure established during serial learning included remote forward associations as well as adjacent ones. These are depicted in Figure 6.1, where solid lines denote adjacent associations and dashed lines from each element to every following element in the series denote the remote

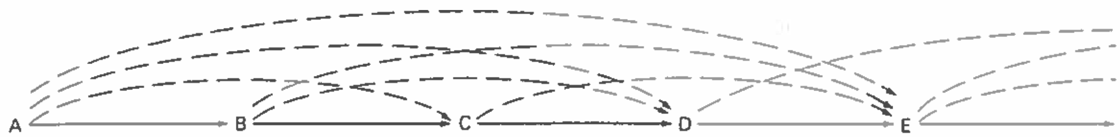

Figure 6.1. Diagram of adjacent and remote associations formed during the learning of a serial order of items.

associations. These remote associations could supposedly be activated at any point in the series, leading to anticipatory errors of decreasing degrees of remoteness. To reach perfect performance, the correct adjacent associations must be strengthened relative to the remote associations. Ebbinghaus believed that the existence of such remote associations (order errors) challenged the doctrine of association by simple contiguity, since stimulus *A* would not be present immediately contiguous with a remote item like *C*, *D*, or *E*. This problem was handled in a later refinement of the theory by Lepley (1934) and Hull (1935). They assumed that each stimulus item left a decaying stimulus trace which endured and was present when later responses in the series were evoked and reinforced. Therefore, associating these later responses *C*, *D*, *E* to the trace of stimulus *A* was a mechanism for producing graded strengths of associations to remote items in the series. The process is depicted in Figure 6.2.

Besides anticipatory errors, evidence offered for the existence of remote associations by Ebbinghaus was that learning a first list resulted in some "savings" toward learning a new list derived from a previously learned one. If the original list is *A*, *B*, *C*, *D*, *E*, . . ., then a first-order derived list would skip every other item as in *A*, *C*, *E*, . . . and a second-order derived list would skip two items as in *A*, *D*, *G*, ... Using himself as the sole subject, Ebbinghaus found that, compared to a scrambled list, derived lists were somewhat easier to learn, and slightly more so the smaller the gap in deriving the remote list. This seems to be consistent with predictions of his doctrine of remote associations.

While Ebbinghaus's doctrine of remote associations led to much research, the ultimate judgment must be that it was incorrect and a misinterpretation of the data. First, the existence of order errors per se need not imply remote associations between items. One could just as well imagine that the items become associated to some internal representations of serial positions (for

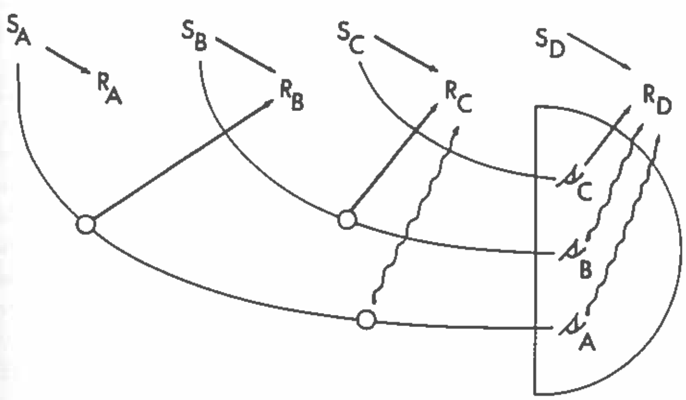

Figure 6.2. Stimulus-response representation of the series *A*, *B*, *C*, *D*, ... Presentation of each stimulus leads to its pronunciation. Traces of the stimuli (small *s*'s) perseverate and become associated (wiggly lines) to more remote responses in the series.

example, the ordinal numbers "first is *A*, second is *B*, third is *C*," and so on), and that order errors reflect stimulus generalization of responses among similar ordinal stimuli, as discussed in Chapter 2 regarding the spread of effect. This would explain the gradient of remote errors, as well as the fact that serial recall errors are somewhat likely to be backward (e.g., saying *B* after *E*) as well as forward. Second, the derived list method seems to have produced slight positive transfer for Ebbinghaus because he knew or could detect its principle of construction (e.g., skip every other item) and so could directly utilize his knowledge of the initial series to guess correctly on the derived list. Experiments by Slamecka (1964) that used derived lists of variable spacing between items avoided this problem, and his subjects showed no savings whatsoever. More seriously, Young, Hakes, and Hicks (1965) found that derived lists in fact created conditions of *negative transfer,* since in learning a derived list like *A, C, E,* . . ., the adjacent associations *A-B, C-D, E-F* acquired in the original list should be aroused, and will compete and interfere with the acquisition and performance of the new correct associates. Therefore, if familiarization with the items per se can be controlled (e.g., by using familiar words), then the derived list condition should produce negative transfer. As proof of this conjecture, Young and his colleagues indeed found considerable negative transfer (slower learning) of a first-order derived serial list compared to control subjects learning these items for the first time. Therefore, the derived-list methodology has fallen in disrepute, as has the doctrine of remote associations.

In later analyses of serial learning, researchers have distinguished between *item learning* versus *order learning.* The first refers to learners' ability to retrieve the items as unitary responses from memory, whereas the second refers to their ability to place the items retrieved into the correct serial order. Factors which affect item availability, such as the meaningfulness or familiarity, will affect item recall and will thus affect serial recall indirectly. In more recent studies of serial learning, this item-learning component is often circumvented by using familiar responses such as letters or digits; subjects may even know the exact item-set, so their only task on each trial is to remember a specified order of the items.

Serial Position Curve

One of the interesting facts about serial learning is that the ease of learning an item depends upon its position in the serial list. Figure 6.3 shows a characteristic serial position error curve; items at the beginning and end of the list are learned fastest, while items just beyond the middle are the most difficult. The relative shape of such serial position error curves seems to be invariant over changes in many variables that affect overall learning rate on the list, e.g., mean-

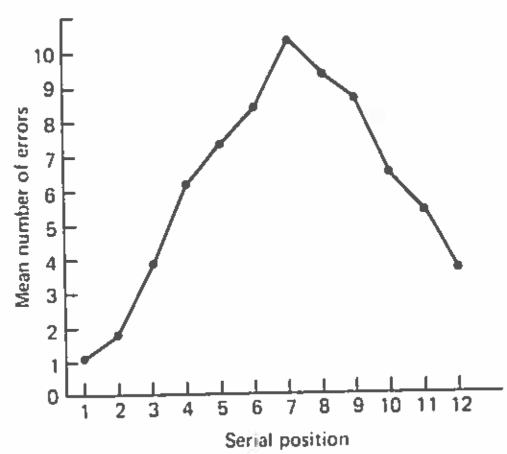

Figure 6.3. Errors at each serial position in the learning of a 12-item serial list of nonsense syllables to mastery. Errors include failures to respond. (From Hovland, 1938.)

ingfulness of the items, presentation rate, intertrial interval (McCrary & Hunter, 1953). The absolute number of errors varies with such factors, but the relative percentage of all errors that are attributable to errors at each position yields much the same curve. Because of the appealing simplicity of this invariance, many theorists have tried their hand at explaining it. Some of these efforts will be touched on in the subsequent discussion.

An early theory of the serial position curve was that of Lepley (1934) and Hull (1935) which made heavy use of the idea of remote associations and the acquisition of inhibitory connections to suppress these remote errors. These inhibitory factors were assumed to pile up most in suppressing responses in the middle of the list, and so most errors should occur at the middle positions. By some strange logic, it was assumed that inhibitory connections spanning adjacent positions (thus, the inhibition not to give *E* to *B* spans the pair *C-D*) caused errors at that interior position. The details are unimportant here since the basic premises of the hypothesis (i.e., remote associations) have been discredited as has the theory constructed on that basis.

An interesting second theory of the serial position effect was proposed independently by Jensen (1962) and Feigenbaum and Simon (1962). Jensen's exposition will be outlined here. It applies to the situation in which a subject is exposed to repeated trials through the same serial list, so that the end of the list is soon followed by the beginning of the list for the next trial. Jensen assumes that the items learned first (or best) are usually the ones to which the subject first attends, or in fact the first item or two in the list. These first learned items then serve as an "anchor point" for learning the remainder of the list. It is assumed that the subject learns most readily by attaching new items to previously learned items. This implies that items are learned around the anchor points both in a forward and backward direction. The learning spreads out around the first one or two items in either direction. The order of learning for the data in Figure 6.3 can be predicted by folding the 12-item series around the anchor point of the first two items as follows (start in the middle!):

Serial Position:	8,	9,	10,	11,	12,	1,	2,	3,	4,	5,	6,	7
Order of Learning:	11,	9,	7,	5,	3,	1,	2,	4,	6,	8,	10,	12

The predicted order of learning is items 1, 2, 12, 3, 11, 4, 10, 5, 9, 6, 8, 7. The rule is to start with the first two items and then alternate successive items from the end of the series and away from the beginning of the series. This notion for predicting serial position difficulty appears quite valid; the average correlation between rank of predicted position difficulty and obtained errors was about .97 over some 70 serial position error curves that Jensen collated from the experimental literature. For example, for the Hovland data shown in Figure 6.3, Jensen's rule predicts the rank of error scores with only one slight misordering (items 6 and 9 are reversed from predicted). The fit of the theory is about as high as the reliability of the serial position curves obtained in different studies.

A problem with Jensen's approach is that it is basically a rule of thumb that describes serial position curves, but the underlying mechanism on which it is based—that is, attaching items to expanding anchor points—seems somewhat implausible. Furthermore, some tasks such as immediate recall of a series heard once (the familiar memory span test) yield neat serial position error curves like those in Figure 6.3, yet it is difficult to see how the anchor point theory really applies to a one-trial task (e.g., learning items at the end of the list as "near" the anchor point of the first item). Because the basic learning mechanism is somewhat vague and implausible, Jensen's theory has not been widely accepted despite its undoubted ability to predict relative serial position error curves.

Serial-position distinctiveness. An attractive recent theory of the serial position curve considers it to be a special case of the differing distinctiveness of positions along any ordered stimulus series. This theory say that some representation of serial position is an important stimulus component for the items of a serial list, and that the ends of the list are more discriminable or distinctive, and therefore better stimuli, than the interior positions of the list. Murdock (1960), Ebenholtz (1972), and G. H. Bower (1971) have articulated this theory and reviewed many studies showing that serial position curves arise whenever the subject must learn to assign different responses to stimuli that vary along a single dimension. An example would be learning to assign letters or names consistently to different lengths of lines, or shades of gray, or pitches of tones, or spatial locations of a dot, or intervals of elapsed time. In the typical experiment, these stimuli would be presented singly in random order, as in the paired associate paradigm. In each case, fewest errors would occur for assignments to the end stimuli and most errors to the interior ones. Bower (1971) pointed out that such results are predictable by assuming equal generalization gradients of responses associated to each stimulus position (see Figure 6.4); more errors occur in the middle of the series because generalized responses can be intruded there from similar stimuli on either side of the target stimuli. In contrast, the correct response is relatively dominant to the end stimuli since intrusion errors can generalize only from one side. This can be seen by simply adding up generalization tendencies for error responses at each stimulus in Figure 6.4. This theory is much like Spence's earlier theory regarding the difficulty of middle-size discriminations (see Figure 5.7, p. 116). Such

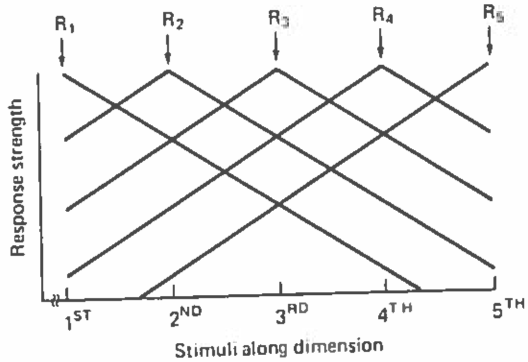

FIGURE 6.4. Strength of each response, R_i, associated to its corresponding stimulus, S_i, with generalization to similar stimuli spaced along the stimulus dimension.

theories handle the rote-serial learning results by supposing that the subject sets up something like "serial position markers" (e.g., first, second, . . .) to which she associates the successive list items, but the markers are more or less distinct and so the responses generalize between nearby position markers. Ebenholtz (1972) reports many experiments demonstrating the validity of positional learning as a dominant mode in serial learning.

To take matters one step further, Ebenholtz and Bower have suggested that the serial positional markers may be abstract and can be transferred to different sets of linearly ordered stimuli, either within the same sensory dimension or to a different dimension. For example, after learning to assign a set of nonsense syllables to lines of differing lengths, Ebenholtz's subjects showed positive transfer when the same responses were transferred and assigned in the *same order* as before to a set of gray patches varying in brightness. That is, the response learned to the shortest line was assigned to the darkest patch, and the remaining responses were kept ordered as learned previously. This transfer, done with college student subjects, may be simply explained by supposing that subjects are converting the stimuli of the set into some abstract codes such as "least, next to least, . . . , middle, . . . , next to most, and most," and then associating specific responses to these abstract codes. In any event, such a theory seems to have moved rather far afield from Ebbinghaus's original theory of remote associations.

PAIRED-ASSOCIATE LEARNING

In paired-associate learning (PAL), an explicit stimulus is provided for each response term. Normally a number of pairs are learned concurrently by the subject using the anticipation method—that is, the stimulus is presented, the subject responds, feedback is given. Paired-associate learning became increasingly popular because of its obvious face validity for the stimulus-response associationism that dominated human learning research in the years following 1940. The simplicity of PAL is only apparent, however, since the results rapidly become complex.

Following the program of constituent analysis, researchers in verbal learning have divided PAL into three component processes: learning to discriminate among the stimuli, learning the responses as units, and learning to associate the correct response unit to each stimulus. Modern theories of PAL deal with all three processes. An experiment by McGuire (1961) was exceptionally clear in illustrating these processes and, in addition, permits an assessment of the contribution of each factor to the performance of the complex behavior exhibited in PAL. Here we will touch on material relating to these component processes.

Discrimination Learning

Paired-associate learning clearly involves stimulus discrimination learning. In a classic paper, Eleanor Gibson (1940) recognized this and systematically applied to

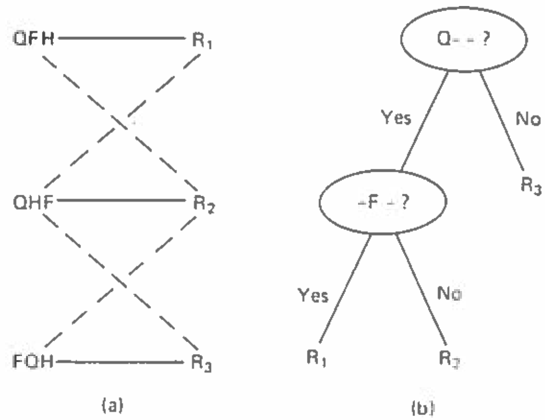

Figure 6.5. (A) Paired associate learning according to Gibson's habit-and-generalization theory and (B), according to EPAM's discrimination network. In A, solid lines indicate correct associations, whereas dashed lines indicate a few generalized error tendencies.

PAL Hull's theory of discrimination learning. This theory assumes that habit strength accrues between each stimulus and its correct response during each reinforced trial, that a response associated to one stimulus may generalize to other similar stimuli causing confusion errors or blocking of the correct response by the generalized one. Furthermore, because generalized responses are nonreinforced, the person will develop inhibition and suppress specific generalized errors at each stimulus to which they occurred. On this view, errors to a given stimulus cease when the correct response has a net strength (habit minus inhibition) exceeding the net generalized strength of each of the competing responses.

While the theory is reasonably complex, it does predict a number of well-known facts about PAL. First, there clearly is stimulus generalization during or after PAL. Thus, a response associated to a 3-letter stimulus like *Q F H* will be found to occur to some extent to *Q E H* or *Q F P* or *Q-H*. Second, PAL proceeds more slowly the greater the similarity among the stimuli in the list. Thus, it will take longer to learn names to the stimulus set *QFH, QHF,* and *FQH,* which overlap in elements than to *ABC, DEF,* and *GHI,* which do not overlap at all. Third, in later work, Gibson extended the theory by assuming that subjects could be taught discriminating features of the stimuli so that positive transfer would be obtained when the same stimuli were associated with a second set of different responses. An experiment by Goss (1953) found such an effect of "predifferentiation" experiences in which subjects compared stimuli and noticed in what ways they differed; later PAL of new responses to these stimuli was more rapid than for control subjects previously exposed to different stimuli.

Later research has accepted the general facts about similarity effects on PAL but not dealt kindly with Gibson's particular formulation of the role of inhibitory factors in the process. Some later theories of stimulus discrimination in PAL proceed in a radically different manner from Gibson's early attempt. Thus, Simon and Feigenbaum (1964) developed the EPAM model which simulates PAL by developing a discrimination network or sorting tree during its experience in PAL learning (see panel B of Figure 6.5). Each node in the tree asks a test question of the stimulus ("Is the first letter a Q?") and is followed by two branches (to lower nodes) that are taken, depending upon the outcome of the test. Responses are stored at the bottom of the sorting tree and are retrieved and output when a stimulus is sorted to that terminal. Two stimuli are confused if they are sorted to the same terminal of the tree; confused stimuli can be distinguished by creating a new test node based on a distinguishing feature of one of them and entering this into the tree. Details of the EPAM model and its successor, SAL (Hintzman, 1968), are presented in Chapter 12 on information-processing theories. Suffice it to say that they account for the types of data marshalled for Gibson's theory, and then explain considerably more results besides.

Response Learning

Response learning refers to the acquiring of the nominal "responses" in the experiment as available units of memory. The nominal response in a PAL experiment may itself be a novel chain of elements such as a nonsense syllable or 3-digit number, and these must be learned as units. Typically, a majority of errors in such nonsense syllable experiments are failures to respond, or incomplete or garbled versions of the appropriate response terms. Thus, in learning the pairs 1-*QHJ*, 2-*QXJ*, and 3-*HJX*, the person may produce 1-*QJX* or 2-*QHJ*, and an error would be recorded. Clearly, the subject in such a task must learn several miniserial lists ("3 then *H* then *J* then *X*") and overcome confusions and interference among them.

The most potent factor controlling the response learning of nonsense syllables is how closely they approximate familiar letter sequences, specifically words. The more wordlike a nonsense trigram appears, or the closer it comes to matching statistics of real English words, the quicker it will be learned and given as a response in PAL. If the subject is prefamiliarized with a set of nonsense syllables, his later PAL will be facilitated when these units are used as responses. These effects of item learning occur not only in paired associates but also in serial learning and in free recall learning. Response learning effects are understandable simply as miniature serial learning tasks that are embedded within the overall task.

Association Formation

Association formation refers to the hooking up of discriminated stimuli with integrated response units. This process has served as a focus for much research. We will not review it here, but will merely point to two research topics surrounding association formation.

Incremental vs. all-or-none learning. A first topic concerns the time course of formation of an association over successive practice trials. If precise measurements could be made of the probability on each trial that a single stimulus evokes the correct response from a subject, what would this curve of response probability look like when plotted over successive practice trials? Would it increase gradually trial by trial or would it consist of one or more discrete jumps from one probability level to a higher one, each level being maintained for several trials? The answer is not obtainable directly since on a single trial for a single S-R pair we observe either a success or an error, but neither gives sufficient information to infer much about the underlying probability learning curve. Consequently, the attacks on this question have been indirect, examining implications of the gradual incremental view as opposed to the all-or-none discontinuous view of learning. Hull and the verbal learning researchers have generally championed the gradual incremental view; Guthrie, Estes, and the Gestalt psychologists have generally championed the all-or-none discontinuous view. For a review, see Restle (1965) or Bower (1967b). This issue will be taken up in more detail in Chapter 8 on stimulus sampling theory.

Mediators. A second issue surrounding association formation is the role of *mediators* in constructing a bridge from the nominal stimulus to the nominal response. A mediator is some bit of knowledge or some preexisting association that the subject comes up with to help her learn the *A-B* association. For example, to learn the pair *RZL-CAT* the subject might say she used the chain *R*-rat-cat or *Z*-hissing-*CAT* or *L*-leopard-*CAT*. Such mediators (also called

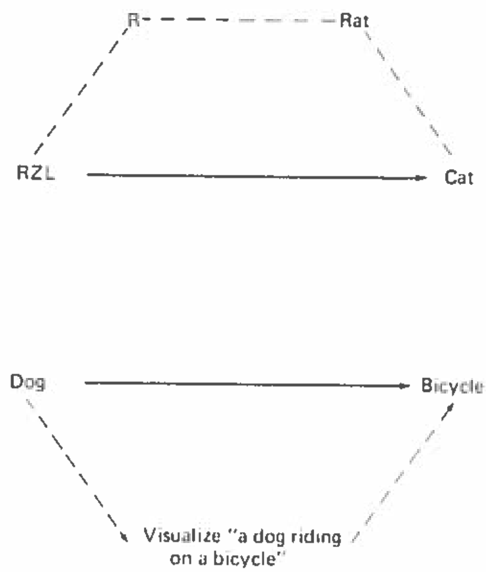

Figure 6.6. Two examples of mediators. The top mediator (dashed line) uses a selected letter-word associate to learn the pair *RZL-CAT*. The bottom shows use of an interactive image to associate a pair of nouns.

mnemonics) use a person's knowledge to find a link (see Figure 6.6). We know that pairs for which the subject can think of a mnemonic, or for which one is supplied, will be learned more rapidly than pairs without mnemonics. The learning not only of paired associates but also of single syllables or trigrams depends upon their triggering stray associations to so-called natural language mediators. Thus, nonsense syllables like *LUV* will be remembered as "LOVE with a spelling change," *CAF* as "cafe without the e," and so on. Prytulak (1971) found that nonsense syllables could be scored in terms of the ease of transforming them into familiar words; moreover, this measure of ease of word-making correlated very highly with the recallability of the nonsense syllables in verbal learning tasks. The results suggest that learning subjects convert a nonsense syllable into a "word plus a transformation," store this code in memory, and then decode it at the time of recall. If subjects remember the word but forget the exact nature of the decoding transformation, they may guess among plausible decodings during output. Thus, they might recall *LOV* rather than *LUV*, *CFE* rather than *CAF*, and so on. Prytulak noted a high frequency of these types of errors that would result from decoding a mediating word incorrectly.

These mediators occur presumably because it is easier to assimilate new material into something familiar plus a correction than to learn the novel combinations from scratch. Similarly, when adults learn *RZL-CAT* using the two partly familiar connections *R*-rat and rat-*CAT* (see Figure 6.6), they are telling us that priming, strengthening, and chaining old associations are more efficient for them than is learning the new associations directly fom scratch.

If adults are asked to learn pairs of meaningful words (say, DOG-BICYCLE), they will often make up a meaningful sentence linking the two concepts into a plausible and memorable interaction ("This DOG was riding on a BICYCLE"). Such mediators improve paired associate learning. Action sentences are better mediators than simple conjunctions ("A dog *and* bike are together"). Furthermore, subjects may also visualize imaginary pictures of the interaction, and this visualization can be shown to greatly enhance memory for the pairings. The benefits are somewhat greater when the subjects generate their own sentences or their own images rather than using one supplied by the experimenter (see Bobrow & Bower, 1969). These sentence-generating or image-generating techniques and their results on recall will be discussed later (see Chapter 13). They clearly form the central techniques of the set of mnemonic devices that have been commercialized in courses on memory improvement (Lorayne & Lucas, 1974). Mnemonic devices are strategies for deliberately recoding material-to-be-learned into a form suitable to be associated via familiar concepts, typically with use of visual imagery. The techniques are quite effective (though less so than commercially advertised), and they have been somewhat researched in laboratory settings (Bower, 1970a; Cermak, 1975). Research on imagery and sentence mediators in associative learning has shaded into the theories of cognitive learning, and will be reviewed in Chapter 13. They seem rather far afield from the older conception that paired-associate learning concerned the establishment of S-R connections by methods analogous to classical conditioning. That older view was laid to rest quite some time ago.

This completes our brief review of functionalist analyses of serial learning and paired-associate learning. The next topics to be reviewed concern *transfer of training* and *forgetting* (or retention). Transfer of training refers to the effects of past learning upon the speed of learning some similar task, to which the earlier habits might be transferred. The transfer may be positive, negative, or neutral, and it depends on the amount and type of overlap in the structure of the two tasks. Because they are such rich domains of large effects, studies of the laws of transfer and forgetting in verbal learning have become the major and continuing enterprise of functionalist psychologists. We therefore review these two research areas in more detail, emphasizing historical developments.

STUDIES OF TRANSFER AND FORGETTING

Our historical review begins with a paper by McGeoch (1932), which provides an early functionalist's account of the conditions that affect transfer and the forgetting of verbal materials. McGeoch accepted two major laws of forgetting and transfer. The first is the *law of context*, which asserts that the degree of retention, as measured by performance, is a function of the similarity between the original learning situation and the retention situation. The second is the *law of proactive and retroactive inhibition,* which asserts that retention is a function of activities occurring prior to and subsequent to the original learning. Proactive and retroactive inhibition have been major topics of research for many decades, ever since the problem was first opened up by Müller and Pilzecker (1900).

The paradigm for retroactive inhibition is *A, B, A,* where the learning of material *B* is interpolated between the learning and the retention of material *A*, and interferes with the retention of *A*. The paradigm for proactive inhibition is *B, A, A,* where the learning of *B* prior to the learning of *A* interferes with the later retention of *A*. Both retroactive and proactive interference with learning are readily demonstrable, and the empirical relationships have led to a number of hypotheses. Studying the development of one of these hypotheses about retroactive inhibition will help us to understand not only retroactive inhibition but the manner in which functionalists construct their theories.

One set of problems concerns the *similarity* between the interpolated material and the material originally learned. E. S. Robinson (1927), arguing from some earlier results of his own and of Skaggs (1925), formulated a hypothesis later christened by McGeoch as the *Skaggs-Robinson hypothesis.* Robinson, following Skaggs, proposed relating the amount of retroactive inhibition to the dimension of degree of similarity between the original and the interpolated material or activity.

With the similarity dimension in mind, Robinson argued that the interpolation of identical material (material *B* the same as material *A*) would simply provide additional practice on material *A*, and hence lead to increased retention on the test trials during which retroactive inhibition is usually shown. Because retroactive inhibition with dissimilar materials was already an established fact, the natural conjecture is that starting with identity the amount of retroactive inhibition would increase gradually as dissimilarity was increased. Now, asks Robinson, what is likely to happen at the other end of the scale, as the original material (material *A*) and the interpolated material (material *B*) become *extremely* unlike? Presumably retroactive inhibition represents some sort of interference based on similarity between the original and interpolated activity. If there is very little similarity, there should be very little retroactive inhibition. Putting all these considerations together, it is reasonable to expect a maximum of retroactive inhibition at some intermediate point of similarity between materials *A* and *B*. Robinson formulated the whole generalization in words as follows: "As similarity between interpolation and original memorization is reduced from near identity, retention falls away to a minimum and then rises again, but with decreasing similarity it never reaches the level of obtaining with maximum similarity." He expressed this graphically as shown in Figure 6.7.

His own experimental test of the generalization was very simple. By the memory span method he studied the recall of the first four of a series of eight consonants as this recall was interfered with by the last four of the consonants. That is, the first four were considered to be material *A*, the last four material *B*, and the similarity and dissimilarity of materials *A* and *B* were controlled. Similarity was defined in terms of common letters in the two halves of the series. Maximum similarity means that the second four consonants were exactly the same in the same order as the first four; maximum dissimilarity meant that all the last four differed from the first four.

The results confirmed the hypothesis only partially. Starting with near-identity, retroactive inhibition increased as the interpolated material became increasingly dissimilar. This was part of the conjecture.

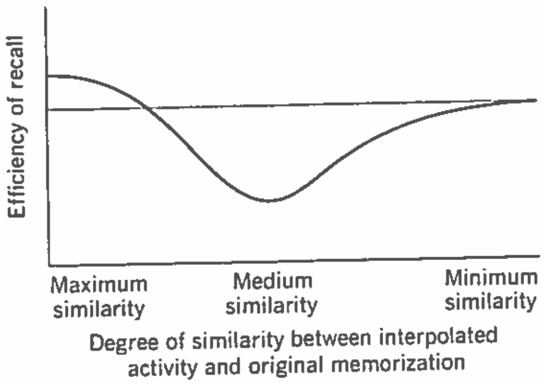

Figure 6.7. Similarity as a factor in retroactive inhibition. The curve is intended to show that retroactive inhibition bears a quantitative relationship to the degree of similarity between the interpolated activity and the material originally memorized. With maximum similarity, the interpolated activity provides positive transfer, hence increases the efficiency of recall. Maximum interference with recall is predicted to fall at some intermediate value of similarity. (From E. S. Robinson, 1927.)

But the decrease in the amount of retroaction (increase in recall) with maximum unlikeness was not found. In fact, with the materials totally dissimilar, retroactive inhibition was at a maximum.

Later investigators had no better luck than Robinson did in confirming his transfer curve. In fact, later analyses of verbal learning tasks began to uncover several distinct *sources* of "similarity" between two tasks as well as several distinct *kinds* of similarity. For example, in paired associate learning, the two successive lists of pairs may be schematized as S_1-R_1, then S_2-R_2; at least two sources of similarity are stimulus similarity (of S_2 to S_1) and response similarity (of R_2 to R_1). Second, the kinds of similarity may be in terms of overlap of common elements in nonsense syllables (e.g., *VAX* and *VAS* are said to be "formally similar") or in terms of semantic or associative meaning of two words (e.g., *elated* is somewhat similar to *high*, less to *low*, and the opposite of *sad*). Recognition of these complexities spelled the demise of the Skaggs-Robinson hypothesis. In reviewing the history of studies of transfer and retroaction, Postman (1971, p. 1083) has this to say:

> In retrospect it becomes apparent that the Skaggs-Robinson hypothesis failed because it was essentially a nonanalytic formulation, which did not specify the locus of intertask similarity. The hypothesis lapsed into disuse as the analysis of similarity relations in retroaction, as in transfer, shifted to the investigation of stimulus and response functions.

Osgood (1949) proposed a more complex formulation of the relationships involved in transfer. His proposed mapping of similarity effects in paired associate transfer is diagrammed in the three-dimensional surface shown in Figure 6.8. What Osgood's surface states is that the amount of transfer in positive or negative directions is a function of shifts in similarity of *both* the stimulus conditions *and* the response required. Shifts in stimulus similarity are from front to back as noted on the right-hand margin, moving from identical stimuli (S_I) through similar stimuli (S_S) to neutral stimuli (S_N) that are far distant on a generalization gradient. Shifts in response similarity are represented from left to right, as noted along the back margin, with identical responses (R_I) being at the left, and moving progressively through similar responses (R_S), neutral responses (R_N), partially opposite responses (R_O) to directly antagonistic responses (R_A). For meaningful verbal materials, the "antagonistic response" was defined as the antonym, a word having the opposite meaning (e.g., *elated–sad*).

The best way to read the diagram is to read its edges first. The rear edge indicates that stimuli bearing *no* resemblance to those used in original learning do not result in any transfer effect, positive or negative, regardless of the degree of resemblance between the required responses and responses that have been used in earlier ex-

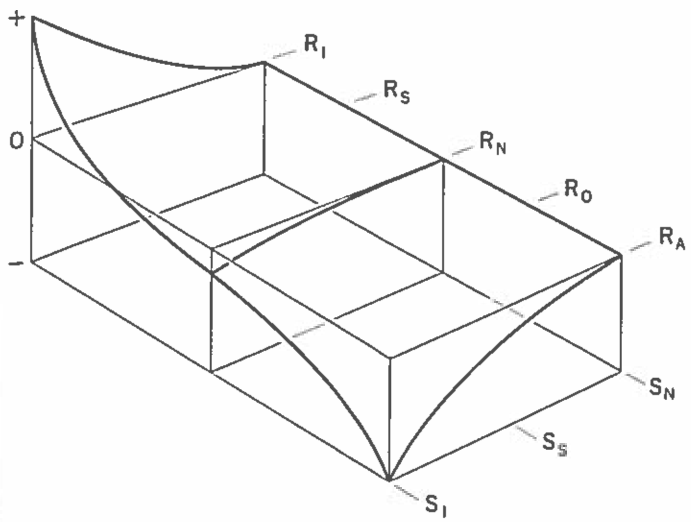

Figure 6.8. Osgood's transfer and retroaction surface. Vertical dimension, amount of transfer (+) or interference (−), with neutral zone represented by a plane (0). Left to right, amount of shift in response similarity between original task and new task, from identity (R_I) to antagonism (R_A). Front to rear, amount of shift in stimulus similarity between original and new task, from identity (S_I) to neutrality (S_N). (From Osgood, 1949.)

periments. The front edge indicates that with *identical* stimuli there will be maximum positive transfer with *identical* responses (for this is merely overlearning); whereas with directly *antagonistic* responses (opposite in meaning) there will be maximum interference, for the earlier responses will have to be completely unlearned or overcome. The left edge indicates that, with identical responses, shifts in stimulus similarity from identity to neutrality will result in decreasing transfer, but no interference in new learning. The right edge indicates that, for antagonistic responses, shifts in stimulus similarity from identity to neutrality will produce decreasing interference, but no positive transfer. The diagram is a surface, and yields a curve wherever it is cut by a vertical plane.

Although Osgood's surface represented an important systematic attempt to integrate a large range of transfer and retroaction phenomena, it soon became clear that it, too, was inadequate for a number of reasons. First, there never was any firm evidence that antagonistic responses were associated with more negative transfer than were unrelated responses, even in Osgood's own data. Second, although the verbal learning data give evidence of differences in transfer between identical, similar, and unrelated stimuli (or responses), they have not shown a *continuous gradient* of effects as similarity is varied over the intermediate range (see Postman, 1971, p. 1054). This graded effect, of course, is implied by the smooth curves drawn. Third, the surface implies that transfer will always be zero when unrelated stimuli are used in successive tasks (see the rear edge of the surface in Figure 6.8). However, subsequent work has shown that this arrangement can produce positive transfer when the response term itself is novel and requires much learning. For example, in order for the subject to recite nonsense syllables as responses to neutral stimuli requires that the syllables be learned as integrated response units per se. Thus, a transfer design like *A-B, C-B* (dissimilar stimuli, identical responses) avoids the necessity of learning the second-list responses, and so can produce positive transfer on this account. (To keep straight the various transfer designs, refer to the illustrations in Table 6.1).

As a fourth complication with Osgood's surface, experiments using a design like *A-B, B-D* (the response in the first list serves as the stimulus in the second list) made it apparent that a so-called *backward* associ-

TABLE 6.1. **Paired associate materials which illustrate different relations between an originally learned list (A–B) on the far left and an interpolated list to be learned. In each pair, the cue word used to prompt recall is the left-hand one.**

A → *B*	*A* → *D*	*C* → *B*	*C* → *D*	*B* → *D*	*A* → *Br*
dog → pin	dog → shoe	card → pin	card → shoe	pin → shoe	dog → sky
cup → mat	cup → tree	book → mat	book → tree	mat → tree	cup → pin
desk → sky	desk → rug	car → sky	car → rug	sky → rug	desk → mat

ation was being established (from *B* to *A*) at the same time that the person was learning the forward association (from *A* to *B*). This backward association then intrudes and causes negative transfer in the *A-B*, *B-D* paradigm, although it causes positive transfer in the *A-B*, *B-A* design (one just reverses which items serve as cues and which as responses). As a fifth problem, the similarity relations treated in Osgood's surface deal with relations among individual items (pairs) across successive lists, and not with the overall structural relations between successive lists. But it is known that the greatest degree of negative transfer in verbal learning occurs when the stimuli and responses of the first list are simply *re-paired* in new ways to compose the items for the second list. In symbolic notation, this is denoted as the *A-B*, *A-Br* paradigm (see last column of Table 6.1). Within the framework of Osgood's surface, this *A-Br* condition can only be represented as *A-B*, *A-D* with identical stimuli and different responses. Nevertheless, it is known that the *A-Br* paradigm produces much more negative transfer (largely due to competing backward associations) than does the *A-D* design.

As a sixth and final complication, Osgood's surface implies that negative transfer in the rate of learning a second list would be perfectly correlated with the amount of forgetting (retroactive interference) of the first list caused by the subject's learning of the second. While these two measures are frequently correlated, some discrepant cases are now well known. One of these discrepancies is that whereas *A-B*, *C-D* (unrelated stimuli and responses) serves as the baseline for defining zero transfer in second-list learning, it is clear that *C-D* interpolation causes extensive forgetting of *A-B* (or nonspecific interference) (Newton & Wickens, 1956). The forgetting of *A-B* by such *C-D* subjects is quite large compared to that of control subjects who learn *A-B* and then simply rest for an appropriate interval before a retention test. Adopting the *C-D* baseline, then, the *C-D* condition produces zero negative transfer but considerable retroactive interference. This observation, in fact, has been one of the reasons for recent doubts regarding the existence of retroactive interference specific to particular paired-associate stimuli (see Postman & Stark, 1969).

Quite clearly, transfer in paired associates is not a unitary process, but is rather composed of a number of distinct components which come into play during initial learning and transfer testing. As stimulus or response similarity is varied, different aspects or components of the transfer task

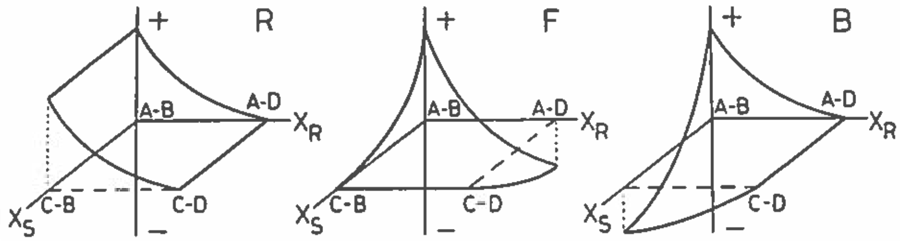

Figure 6.9. Component surfaces for response availability (R), forward associations (F), and backward associations (B). Degree of stimulus similarity is spaced along the X_S axis, and degree of response similarity along the X_R axis. The point of origin represents identity of both the stimulus and the response. Direction and degree of transfer are indicated along the vertical axis. (From E. Martin, 1965.)

will vary, although the net effort on performance may be unclear because different components contribute positive or negative effects which may cancel or nullify one another. Martin (1965) has published a more informed attempt to specify transfer surfaces for three different components carried over in transfer; this is shown in Figure 6.9. The three components of transfer considered are response learning (R, in the left-hand panel), forward associations (F, in the middle panel), and backward associations (B, in the right-hand panel). For response learning, the degree of positive transfer is high for identical responses, decreasing to zero for more dissimilar responses, and is independent, of course, of stimulus similarity. The surface for forward associations (F) is the same as Osgood's except that Martin suggests that the response continuum should extend only to unrelated responses (*A-B, A-D*); although antonyms are opposites in meaning they are associatively close to one another and this relation, more than their meaning opposition, dominates so as to produce slightly positive transfer when antonymic responses are used (see Postman, 1971). The results for backward associations (B) are rather symmetrical to those for forward associations except for an interchange of stimulus and response axes. In particular, interpolation of *C-B* following *A-B* learning produces maximal loss of the backward association from *B* to *A*, as indicated in the subject's failure to recall the first-list stimulus *A* when cued with the response term *B*.

Martin's hypothesis does not indicate specifically how these several factors will combine to determine the net transfer effect. It is clear, too, that even further factors have been identified which can influence transfer and forgetting, and these features are not depicted in Martin's surfaces. Thus, for example, a number of experiments have shown that there are some positive benefits from prior attentive exposure to the stimulus or the response terms since this serves to "predifferentiate" these items prior to the commencement of the criterion paired-associate task. To the extent that the subject has already learned to identify and discriminate among the stimuli, it is that much easier for him to continue discriminating while attaching particular responses to the predifferentiated stimuli. Although the normal *A-B, A-D* paradigm involves this predifferentiation factor, the competing response factor seems to override it and to determine net negative transfer in most circumstances.

Another factor missing in this transfer surface is what has been called *list differentiation,* the ability of the subject to identify the list membership of responses which he recalls. Thus during *A-D* learning following *A-B*, the person may intrude the *B* response, although he knows both responses, because he confuses the two lists. Or he may think of but withhold response *D* because he erroneously identifies it as coming from the first list. In like manner, following interpolation of *A-D*, if the person is asked to recall the first-list response, he may withhold *B* or intrude *D* if he is unable to distinguish which list the various responses occurred in. List differentiation as a process is closely analogous to remembering the time and context in which events occur, a topic of considerable interest in its own right (Anderson & Bower, 1972a; Hinrichs, 1970; G. H. Bower, 1972d). List identification will clearly be a factor in transfer, since confusions about list membership will increase with similarity of the stimuli or responses. Thus, for example, if the list-2 responses are digits whereas the list-1 responses are nouns, there would probably be perfect list differentiation and the subject would almost never intrude a digit response while trying to recall a list-1 noun.

Martin's hypothesis does encompass the massive negative transfer produced by the re-paired *A-Br* transfer task, since Figure 6.9 shows negative transfer for the forward association (at *A-D* in the middle panel) and negative transfer for the backward association (at *C-B* in the right-hand panel). What it fails to show is that the *A-B, C-D* paradigm produces zero negative transfer but nonetheless appreciable forgetting relative to a rest control condition. This is an issue we will touch upon later.

Let us retrace the evolution of these hypotheses (or guessed-at "empirical generalizations") regarding similarity effects in retroaction and transfer. First we have the early experiments demonstrating retroactive interference, and some that demonstrate the possible role of similarity as a factor. Next we have Robinson's somewhat crude dimensional hypothesis, leading to a series of experiments which reveal multiple sources and kinds of intertask similarities, requiring a generalization more complex than the Skaggs-Robinson hypothesis. At the same time these investigations of forgetting were occurring, a series of related experiments on transfer were being conducted, but as though the two phenomena had little in common. Presently Osgood offered his new synthesis, covering the data that had accumulated since Robinson's hypothesis was announced, incorporating in one stroke the results on transfer and on retroactive interference. But a series of analytical studies revealed that the Osgood surface was too simplistic, that there are even more components or independent factors involved in transfer. These more detailed analyses led to Martin's proposed *component transfer* surfaces for three of the important bits of learning that are now known to be carried over in transfer. Although Martin's proposals are the most adequate integrative summaries yet seen, it is clear that several isolable factors have still been ignored, and the weighting of the magnitude of the independent factors and their interactions in determining net transfer has been left for future specification. This sort of succession of hypothesized generalizations, with an interplay between data, analytical criticism, and theory, seeking a deeper analysis but a more revealing integration, is exactly what might be expected in a maturing functional analysis. This history also illustrates some of the potential frustrations of a functionalist approach; upon closer analysis, more and more variables or independent factors are discovered to influence the behavior under investigation, and the possibility of strong interactions between one variable and the functions obtained for other variables is a likely prospect. For example, the relating transfer function to response similarity probably varies in quantitative shape depending on whether one is dealing with "formal similarity" (overlapping elements of nonsense materials) or with "meaningful similarity" (e.g., synonymous words); and it is unclear yet whether the latter variable should be anchored to conceptual overlap of dictionary definitions of two words or to overlap of the associative hierarchies elicited by each word (see, for example, Deese, 1965).

The complexities of the behavioral phenomenon of transfer make some investigators despair (e.g., to say "nature couldn't be *that* complicated"), give it up as a poorly formulated scientific question, and move on to work on other issues in psychology with greater prospects of quick progress. The dyed-in-the-wool functionalist would argue that he was simply doing the yeoman's work of elucidating a phenomenon, and claim that no one is ever guaranteed simplicity in his findings and that, although the subproblems shift with the maturity of his analysis, the overriding phenomenon with which the area began (viz., transfer of training) is undeniably a central problem for all of learning theory. If it truly is a central issue, then, the functionalist would claim, it must be studied, analyzed, and understood with the only experimental and conceptual tools we presently have available. We must analyze and understand complexity because "that's the way the world is."

Analysis of Forgetting

The foregoing account of research on negative transfer and retroactive interference was carried out at the level of empirical description and generalization without much interest in theory. However, by far the most significant portion of the research of modern functionalists concerns the analysis of theoretical mechanisms of forgetting. The character of the functionalists' theoretical approach can perhaps be best appreciated by tracing the evolution of their ideas regarding forgetting. Most of this research has been done on verbal learning experiments with human adults. Of course, animals forget too, even simple conditioned responses, and the study of forgetting in animals has become a major area of study (see Honig & James, 1971; Spear, 1978). To reduce the difficulty of the task of understanding, recent experiments on forgetting have concentrated on standard verbal learning situations to yield the main evidence.

If one asks the layman why he forgets things, he has a ready answer: he forgets things because he hasn't used them, or thought of them, for some time. He has forgotten the Spanish he learned in high school because he hasn't used it for the past ten years. But he remembers things he continues to use, such as the names of his friends.

The problem with this popular account is that it does not satisfy the scientists' curiosity regarding causal mechanisms. Lapse of time is not itself a causal variable, although causal events happen in time. If I leave an iron hammer outside, it will progressively rust with time. But it is not the lapse of time that rusts the hammer. Rather, it is the reaction of chemical oxidation that occurs in time.

We can give the layman's proposal a more neurological *sound* (if not sense) as follows: each learning experience establishes a neurological trace whose integrity is gradually obliterated by random neuronal noise that occurs at a fixed rate, eroding away the retrievability of the memory trace as the retention interval increases. Does this formulation buy us anything? The answer is "not really." Unless much more is added regarding relevant variables and their influence on the hypothetical process (and forgetting), the new proposal is worse than vacuous; it is dangerous because someone is likely to consider it seriously due to its apparent technical jargon.

A variety of substantive proposals concerning the causes of forgetting have appeared, differing considerably in their scope and the range of variables of which forgetting is said to be a function. For example, Freud supposed that some forgetting results from active repression of certain materials in the unconscious. A critical discussion of this hypothesis along with the conflicting data surrounding it may be found in an earlier edition (Hilgard & Bower, 1975); for a sympathetic reading of that conflicting literature, see Erdelyi and Goldberg (1979). Another conjecture, contributed by Gestalt psychologists and reviewed in Chapter 10, was that memories were multifaceted systems continually undergoing dynamic change, moving toward some better organization (or gestalt). This notion became translated in laboratory experiments into the question of whether a subject's recall of an asymmetric or incomplete figure or line drawing tends to move during a retention interval toward a "good" or "better" gestalt figure. Riley (1963), in his review of this extensive literature, concluded that there was little consistent support for the Gestalt idea. Recall of a figure pattern, more often than not, does tend to move toward cultural stereotypes but such trends as are found turn out more often to be explainable by verbal associations (to the original figure) or proactive interference from prior cultural learning than by Gestalt laws of perceptual organization.

The most serviceable theory of forgetting that has emerged from laboratory experiments is called the *interference* theory. This is closely tied in to the functionalists' analysis of negative transfer and interference. Currently, interference theory has far more adherents, because of more evidence in its favor, than any or all alternative theories of forgetting, so it is fair to call it the current dominating theme of experiments on forgetting. This is an association theory; that is, its basic primitive concept is the notion of an associative bond (functional connection) between two or more elements, the elements being ideas, words, situational stimuli and responses, or whatnot. As indicated earlier in our discussion of transfer paradigms, the conventional notation uses letters *A, B, C* . . . , to represent such elements or items, and the notation *A-B* to represent an associative bond between *A* and *B* established by some past training. It is presumed that these associative bonds can vary widely in their strength depending on the amount of practice. The experimental situation that best illustrates the theory is paired-associate learning, wherein the subject is taught a set (list) of pairs and then is tested later for retention. The theory applies as well to most other learning tasks, but the paired-associate task makes the expositional mechanics easiest to implement.

Interference Theory

The basic ideas of interference theory were first stated explicitly by McGeoch (1932), but through the succeeding years changes in the theory have gradually occurred. New concepts have been added, unsupported conjectures pruned away, and new experimental methods devised to measure more exactly the relevant dependent variables. The changing character of interference theory may be seen by comparing McGeoch's early statements with Postman's (1961, 1971) later formulations. In what follows, we shall indicate some of the changes and the shifts in emphasis.

The first principle of McGeoch's statement seems an absurd one for a theory of forgetting: it says that forgetting does not occur in an absolute sense. The strength of an association between two items, *A-B,* is established by training, and it is presumed to remain at that level despite disuse of the association. The cause of a measurable retention loss over time is not that the strength of *A-B* decays, but rather that alternative associations, *A-C* or *A-D,* have by some means (to be specified) gained strength in the absence of continued training on *A-B*. Thus, on a retention test, the subject may give *C* or *D* as the associate to

A, so we record a retention loss for the *A-B* association. The *A-B* association has not been lost or forgotten in any absolute sense; it is still there in memory, but *B* has been temporarily displaced, losing out in competition with elements *C* and *D* at the moment of recall.

On the basis of this theory, then, an association once learned is permanently stored, and forgetting is due to declining accessibility, a lessening probability of its retrieval from the storehouse. And this declining accessibility results from competing associations. Such an approach has at least the substrate required to account for the clinically puzzling instances of hypermnesia in which a person demonstrates exceptional recall, or believes his recall is genuine, of experiences from long before. Such heightened recall may occur in manic states, in the hours anticipating some emotionally exciting event (e.g., soldiers about to go into combat), in a hypnotic trance (see Reiff & Scheerer, 1959), or while following a line of free associations when on the psychoanalyst's couch (see Erdelyi & Kleinbard, 1978; Pascal, 1949; Stratton, 1919; Stalnaker & Riddle, 1932).

According to this theory, the *A-B* association may be tested by presenting one of the elements, say *A*, whereupon the subject tries to produce the associated *B*. We may think of *A* as a stimulus term and *B* as a response. As indicated in our earlier review of transfer, this suggests manipulating the degree of similarity of a test stimulus (call it *A′*) to the original training stimulus *A*. The principle of stimulus generalization predicts that *A′* is less likely to activate the *A-B* association in proportion as *A′* is dissimilar to *A*. Moreover, McGeoch suggested that we expand our conception of *A* to include any background contextual stimulation that is present when the *A-B* association is learned. Changes in such contextual stimuli have been found to result in poorer recall (Abernathy, 1940; Falkenberg, 1972; Pan, 1926). Thus, if the subject is tested for recall in a different room than that in which he learned, or with a different type of stimulus-presenting device, or with the material presented on different backgrounds, or when he adopts a different posture, or whatever, his recall is poorer than when, during testing, precisely the original stimulating context is reproduced. Such results seem consistent with the analytic position of interference theory.

Earlier we mentioned that retention loss on a learned *A-B* association results from competition of alternative associations, *A-C*, at the moment of recall. If we ask where these conflicting associations come from, the logical answer is that they (or one similar to them, *A′-C*) come from learning either before or after the *A-B* learning but before the retention test. This analysis has led to the intensive investigation of situations in which the *A-B* and *A-C* learning is explicitly controlled. The two basic paradigms are called retroaction or proaction depending on whether the experimenter's interest is in retention of the first-learned or the second-learned material. These paradigms, together with the appropriate control conditions and some hypothetical recall data, are illustrated in Table 6.2. In the retroaction paradigm, the control group first learns the *A-B* associations, then rests, and later is tested for recall of *B* when given the *A* term. The experimental group learns *A-B*, then learns new pairs *A′-C*, and then tries to recall *B* when given *A*. The retroactive interference index calculated for the hypothetical data is 67 percent. The proaction conditions may be read similarly.

A variety of task variables can be studied in this context, and on the whole the recall results fall in line with what would be expected from interference theory (for reviews, see Slamecka & Ceraso, 1960; Postman, 1971). For example, retroactive interference increases with trials on *A-C* and decreases with trials on *A-B*, whereas proactive interference shows just the opposite functional relations, as expected. Consider just one ex-

TABLE 6.2. Recall results to illustrate forgetting due to retroactive and proactive interference.

	Retroaction		Proaction	
	Experimental	Control	Experimental	Control
List 1	A-B	A-B	A'-C	rest
List 2	A'-C	rest	A-B	A-B
Recall test	A-B	A-B	A-B	A-B
Percent correct recall	20	60	60	80
Effect	$\frac{60-20}{60} = 0.67$		$\frac{80-60}{80} = 0.25$	

ample, namely, the effect of the number of training trials on the *A-B* list (original learning, abbreviated OL) prior to *A-C* learning upon the relative dominance of the *A-B* and *A-C* associations. An experiment by Briggs (1957) illustrates the procedure and results (see Figure 6.10). Four different groups of subjects received 2, 5,

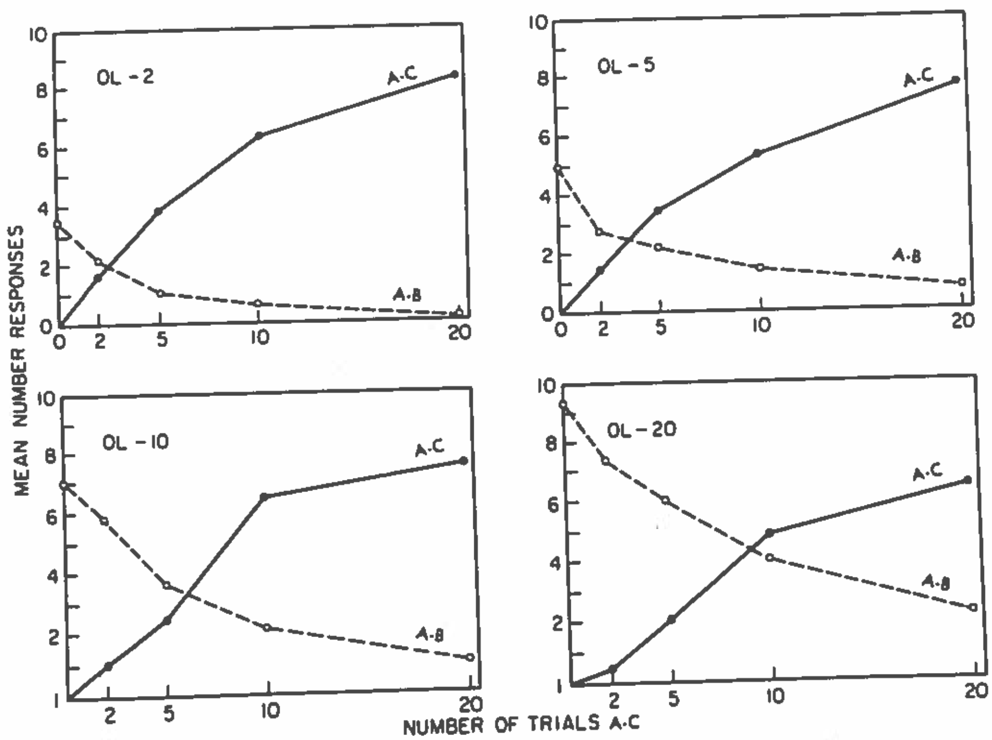

Figure 6.10. Relative response frequencies of the originally learned response (*A-B*) and the newly learned response (*A-C*) during learning of the new response as measured by modified free recall. The four graphs come from four different groups of subjects given 2, 5, 10, or 20 trials of original learning. (From Briggs, 1957.)

10, or 20 trials of *A-B* learning followed by 20 trials of *A-C* learning. The lists were ten paired adjectives. After intervals of 0, 2, 5, 10, and 20 trials of *A-C* learning, each subject received a modified free-recall test with the stimulus terms. On such tests the subjects were instructed to give whatever response first came to mind (including extralist intrusions), and there was no feedback from the experimenter to indicate which response was wanted. Such a test assesses the relative dominance of *A-C* over *A-B*.

The curves in Figure 6.10 provide a graphic description of the frequency of the new *C* response and the old *B* response after varying numbers of *A-B* and *A-C* trials. At the beginning of second-list learning, the frequency of *B* recall depended directly upon the number of OL trials. During the course of *A-C* training, *B* responses decreased in frequency while *C* responses increased to a dominant role. After 20 trials of *A-C*, the amount of OL still exerted some influence, both in terms of a higher *A-B* recall frequency and a lower *A-C* recall frequency. This picture is exactly what one would expect from McGeoch's earlier ideas of response competition, since the modified free-recall test permits only one response.

McGeoch's hypotheses predict a perfect correlation between retention loss of *A-B* and the occurrence on testing of intruding associates, *C* or *D*. This correlation is not always found: on the *A-B* retention test following the *A-C* learning, the subject often is unable to respond with any associate. Two hypotheses were proposed to account for this and both probably have some validity. One notion that we mentioned earlier, proposed by Thune and Underwood (1943), is that the subject can discriminate the list membership (first or second) of associates that come to mind; to the extent that he does this, he will censor and reject response *C* when trying to recall the first-list response, *B*. This is plausible since it is known (Yntema & Trask, 1963) that subjects can judge with fair accuracy which of two events has occurred more recently in the past. Another idea, first expressed by Melton and Irwin (1940), is that during the *A-C* interpolated learning, the first pair *A-B* is unlearned or extinguished. If so, then, when the test occurs soon after the *A-C* learning, *B* is temporarily unavailable as an associate.

The clearest evidence for *unlearning* comes from a recall method first used by Barnes and Underwood (1959). Using the *A-B, A-C* paradigm, the subject was asked on the later test to recall *both* list responses to stimulus *A* and to indicate their list membership. This is a noncompetitive recall situation, and failures are ascribed to unavailability of the responses. The results of Barnes and Underwood are shown in Figure 6.11. This shows that recall of *C* responses increased with trials of *A-C* learning but, more importantly, recall of *B* responses decreased with trials on *A-C*. Thus, as the *A-C* training is extended, the first-list associates become increasingly unavailable, presumably due to unlearning. A variety of follow-up experiments confirmed and extended these results, so that the concept of unlearning was widely accepted.

Postman and Stark (1969) challenged the validity of the associative unlearning concept. They noted that the *A-B, A-C* paradigm produced relatively little negative transfer when they tested the *A-B* pair by multiple-choice *recognition* (that is, recognizing that B_1 but not B_2 was paired with A_1).[1] Furthermore, although this paradigm produced the customarily large *recall* decrement for *A-B*, the forgetting was not

[1] Pair recognition shows strong RI in the *A-B, A-C* paradigm if the interfering response, *C*, is included among the distracting lures on the multiple-choice test for *A-B* (R. C. Anderson & Watts, 1971). However, this RI could be explained by response competition and loss of list differentiation rather than by specific unlearning of *A-B*. Nonetheless, it appears that recognition performance in the Postman and Stark experiment was too high in all conditions to reveal any effect due to associative unlearning.

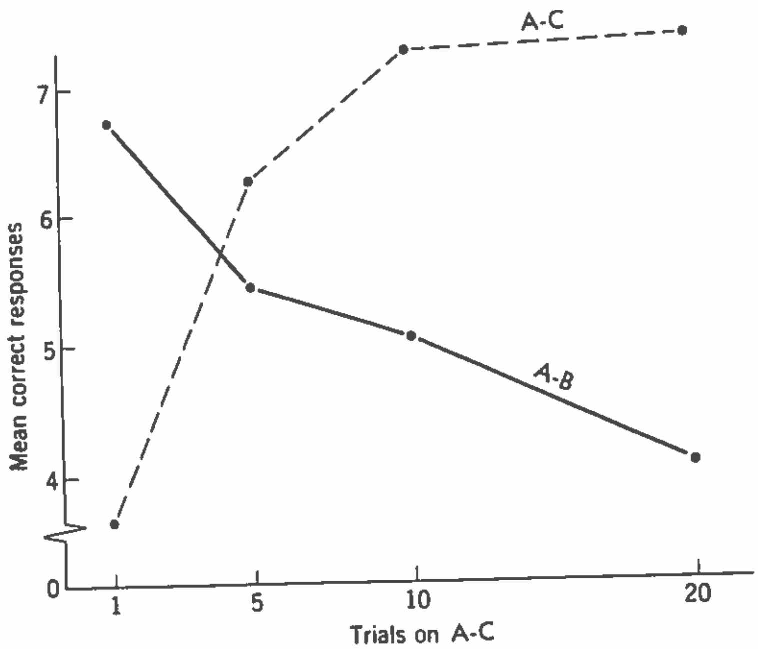

Figure 6.11. Mean number of responses recalled and correctly identified with stimulus and list in the *A-B, A-C* paradigm. Eight is the maximum possible score in each case. (From Barnes and Underwood, 1959.)

very much greater than that produced by an *A-B, C-D* paradigm. For these and other reasons, Postman and Stark suggested that retroactive interference was not being caused by stimulus-specific associative unlearning, but rather was due primarily to a suppression of the entire set of first-list responses, a suppression that develops during second-list learning and persists awhile into a later retention test for *A-B*. On this hypothesis, *C-D* interpolated learning would cause suppression of the *B* response set (as does *A-C* interpolation) and thus these items would be unavailable as responses on the *A-B* recall test. But if the *B* responses were made available as on the pair recognition test, then the person would show that he had not unlearned the *A-B* association.

Evidence for stimulus-specific unlearning was not long in coming following the Postman and Stark challenge. A spate of experiments soon reported demonstrations of this effect (see Birnbaum, 1972; Delprato, 1972; Weaver et al., 1972). A typical one was that by Delprato (1972). He used a "within-list" design in which, across two lists, different items within the list learned by a subject exemplified an *A-B, C-D* relation and other items exemplified an *A-B, A-C* relation. The important point about such a design is that a factor like "suppression of first-list responses" should operate equally on all first-list pairs whether the corresponding second-list item is *C-D* or *A-C*. Therefore, any differences among items in recall of *A-B* could probably be attributed to stimulus-specific learning (i.e., learning *A-C* specifically weakens *A-B* in some absolute sense). Delprato's experiment showed exactly this result, with more forgetting on those specific pairs followed by *A-C* than those followed by *C-D*.

Furthermore, it was possible to prove

stimulus-specific unlearning even with pair-recognition testing (see Merryman, 1971). Apparently, recognition performance in all conditions in the Postman and Stark experiment was too high to provide a sensitive test of specific unlearning. Therefore, we can still retain the idea of stimulus-specific unlearning. However, there does seem to be something to the notion of a general loss of availability of first-list responses due to second-list learning. One may think of this loss of first-list responses as the unlearning of associations between general contextual stimuli and the first-list responses (see McGovern, 1964; Keppel, 1968). It has been proposed that this loss of availability to contextual cues can be studied in the multilist free-recall situation where there are no explicit cues for recall of each item on the list.

Proactive Interference and Spontaneous Recovery

Proactive interference is the decrement in recall of the second-list (*A-C*) material caused by prior learning of the first-list material (*A-B* or *D-B*). Proactive effects are minimal immediately after *A-C* learning, but they increase over a retention interval. It is almost as though the person became confused at the retention test between the two lists he had studied earlier. So one likely explanation of proactive interference is that it involves progressively more confusion between the two lists learned some time before; let us call this the "list-differentiation" idea. It is quite plausible that ability to discriminate between second-list items occurring t hours ago and first-list items occurring $(t + \Delta)$ hours ago will decrease as t becomes larger, a sort of Weber-Fechner law for time discrimination.

But a second explanation has been offered for this increase in proactive interference with an increase in the retention interval following *A-C* learning. This second hypothesis supposes, first, that the original *A-B* associations are extinguished, unlearned, or inhibited during *A-C* learning, and, second, that these *A-B* associates spontaneously recover some of their prior strength over the retention interval. The analogy is to Pavlov's observation (see Chapter 3) that conditioned responses recover during a rest period after a series of extinction trials. Clearly, if the *A-B* associates spontaneously recover, they will compete with *A-C* recall, providing increasing proactive interference as recovery increases over time.

Several lines of evidence support this idea of *A-B* recovery following *A-C* learning. One is an earlier experiment by Briggs (1954), who studied the relative dominance of the *A-B* and *A-C* habits over varying retention intervals using the modified free-recall (MFR) test. In this experiment, subjects learned a first list of 12 paired adjectives (*A-B*) to a criterion of one perfect recitation, rested 24 hours, then learned a second list (*A-C*) to a once-perfect criterion, then received a final MFR test after either 4 minutes or 6, 24, 48, or 72 hours. At various stages during the course of both original and interpolated learning the subjects received an MFR test, in which they were asked to say whatever response first came to mind as they were shown each stimulus term. The results are shown in Figure 6.12. In the left panel is shown the relative frequency of first-list responses (*A-B*) as contrasted to preexperimental associates (*A-E*) from outside the list when tests were given after various levels of first-list performance, the criteria specified in terms of the percentage of pairs correct on the training trial just preceding the MFR test (0, 1/4, 2/4, 3/4, 4/4 of the list). As expected, extralist associates decline, whereas List-1 responses increase. After the 24-hour rest interval, the MFR test revealed some rise in extralist associates and loss of first-list associates (see the 0 point on List 2, the middle panel). Then during List-2 learning to various criteria, List-1 responses and extra-

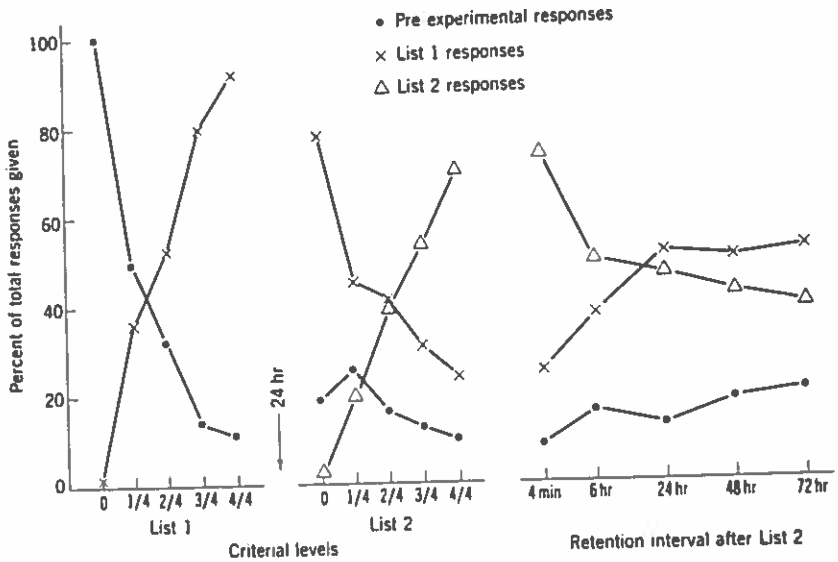

Figure 6.12. The left and center sets of curves are the acquisition and extinction functions when Lists 1 and 2 are learned successively in a retroactive-interference paradigm. Recall as a function of time from the end of List-2 learning is shown in the right-hand set of curves. (Adapted from Briggs, 1954.)

list associates declined, whereas List 2 responses increased.

The data of interest to the spontaneous recovery hypothesis are in the third panel (from separate groups of subjects), showing the relative percentages of *B*, *C*, or *E* responses at differing retention intervals. This graph clearly shows a gradual recovery over time of the *A-B* and *A-E* associates, with a corresponding loss of the most recently learned *A-C* associations. These curves are exactly what would be expected if *A-B* and *A-E* associates were recovering in strength following their unlearning during *A-C* training.

The problem with this interpretation, of course, is that the MFR test is a measure of *relative* response strengths of *B*, *C*, and *E*; Briggs's results could have been produced merely by a greater absolute loss in *A-C* rather than an absolute recovery in *A-B* or *A-E*. The obvious way to proceed is to try to demonstrate absolute recovery of *A-B* in a noncompetitive recall situation, specifically the "modified modified free-recall" (MMFR) tests of the type used by Barnes and Underwood. That is, the subject is to try to recall both *B* and *C* responses when cued with stimulus *A*.

But the evidence for spontaneous recovery in studies of temporal changes using MMFR tests has been equivocal, particularly for longer retention intervals ranging from several hours to several days. However, the possibility remains that a small absolute recovery of *A-B* is being masked by progressively greater recovery of preexperimental associates (*A-E*), which are edited out by the subject in the typical MMFR test. If so, then recovery of *A-B* should be most likely in MMFR tests given at reasonably short intervals after *A-C* learning. In these conditions, positive evidence for absolute recovery of *A-B* has been

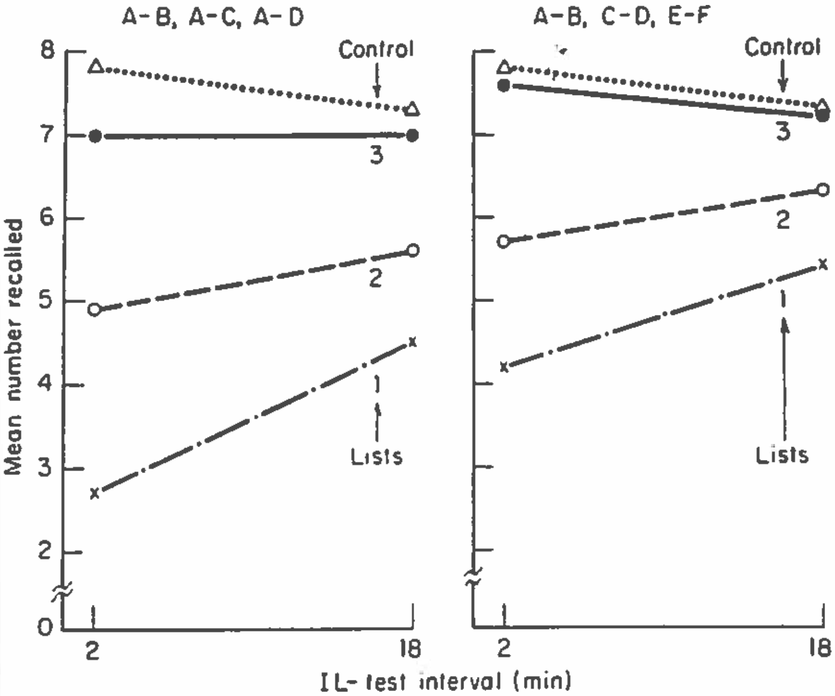

Figure 6.13. Mean number of paired associates recalled from the first, second, and third lists after 2 or 18 minutes' delay for subjects receiving either the *A-B, A-C, A-D* paradigm (left panel) or the *A-B, C-D, E-F* paradigm (right panel). Recalls of control groups who rested for an equivalent time following List-1 learning are also shown. (From Postman, Stark, & Henschel, 1969.)

obtained (see Postman et al., 1968; Postman et al., 1969). Figure 6.13 shows some recovery data from an experiment by Postman, Stark, and Henschel (1969, their Experiment III). The experimental subjects learned three successive lists, having either an *A-B, A-C, A-D* relation or an *A-B, C-D, E-F* relation for different subjects, and then received an MMFR test either 2 or 18 minutes after the final learning trial. Two control groups learned the *A-B* list, then were tested for its recall after a time interval equal to that occupied by the interpolated learning plus 2 or 18 minutes. Figure 6.13 shows significant absolute recovery over the retention interval for both first-list and second-list responses; the third-list response was not unlearned and shows a high level of recall at both retention intervals. Significantly, the amounts of recovery are comparable for the *A-B, A-C* group and the *A-B, C-D* group. This fact suggests that this recovery of early list responses may be due to the dissipation over time of "response-set suppression" rather than to spontaneous recovery of stimulus-specific associative unlearning. That is, during interpolated learning the person may make available the responses being used in that list while selectively suppressing the entire set of responses used in earlier lists. But this suppression dissipates with time, allowing earlier responses to become progressively available for recall in MMFR tests. This issue of the source of recovery, whether of response availability or stimulus-response associations, is one currently under investigation.

The Magnitude of Proactive Interference Effects

One major shift in interference theory that has occurred consists in the powerful role assigned to proactive sources of interference in forgetting. In a major paper, Underwood (1957) employed the proactive idea to clear up what had been a major source of embarrassment to interference theory. Most of the earlier studies of retention had shown rather massive forgetting—about 80 to 90 percent—over 24-hour intervals. The claim that this was due to interference from casual interpolated learning seemed unconvincing since it was difficult to imagine much everyday learning that would interfere with nonsense materials learned in the laboratory. By collating various reports, Underwood determined that those studies reporting massive forgetting had used the same subjects under many list-learning conditions. The more lists a subject had learned, the more she tended to forget the last one when recall of it was measured the next day. Thus, proactive effects presumably accumulated over the lists learned earlier. If a subject learned only a single list of verbal material, then her recall was fairly high—around 75 to 80 percent after 24 hours.

A particularly apt illustration of massive proactive effects is provided by an experiment by Keppel, Postman, and Zavortink (1968). They had five college students learn and recall 36 successive lists (*A-B, C-D* relations) of ten paired associates at 48-hour intervals. Each test session began with a recall test on the prior list learned, followed by the learning of a new list to a criterion of one perfect recitation. The recall percentages are shown in Figure 6.14 plotted in successive blocks of three lists; this shows a dramatic decrease in recall from around 70 percent on the first list to around 5 percent on the last two lists. This illustrates the powerful effects that can be produced by proactive interference. It does not, of course, illuminate the mechanism underlying proactive interference in the *A-B, C-D, E-F* design. Presumably it is loss of availability of the final list's responses; if so, then there should be no cumulative proactive effects demonstrable in pair-recognition tests, or if the response words from successive lists came from distinguishably different (but memorable) semantic categories.

As noted, recall after 24 hours of a well-learned list is about 75 percent, when no interfering lists had been learned. Underwood and Postman (1960) attempted to account for the remaining 25 percent of forgetting observed by appealing to extraexperimental sources of interference. They point out that in learning arbitrary verbal associations or nonsense material in the laboratory, the subject probably has to unlearn the prior verbal habits which he shares with other members of his particular linguistic community. These prior verbal habits may be one of two types—letter sequence associations or unit (word) sequence associations. To give a transparent example, a subject is certain to enter the experiment possessing prior word associations like *table-chair* and *light-dark*. Suppose the learning task requires her to form the new associates *table-dark* and *light-chair*. During a rest interval, spontaneous recovery of the unlearned prior associates will produce a decrement in the probability of her recalling the associations learned in the laboratory. The experiments by Underwood and Postman plus related follow-up studies show some merits of this analysis. New materials that clash with prior verbal habits are forgotten more readily, usually being distorted in the direction of agreement with the prior habits. However, the evidence on this hypothesis has been somewhat conflicting, and it appears that a

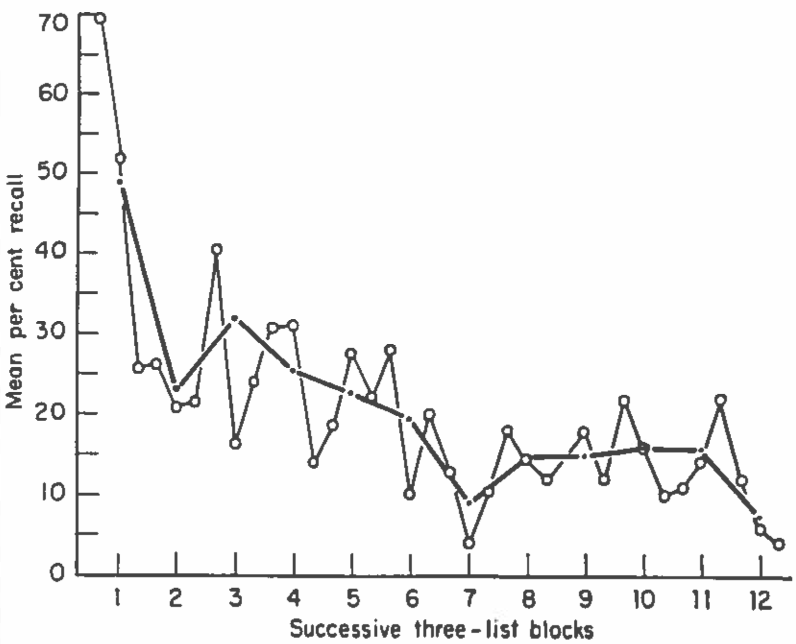

Figure 6.14. Mean percentage of recall of the immediately prior list at a 48-hour interval as a function of the number of prior lists, in blocks of three lists. The heavy black line averages larger blocks of nine lists. (From Keppel, Postman, & Zavortink, 1968.)

variety of complicating factors can enter to obscure the hypothesized relationship.

A powerful demonstration of proactive interference due to prior linguistic habits is to be found in a study by Coleman (1962). He took a 24-word passage of prose from a book and scrambled the words into random order. This order was given to a subject for brief study, and the subject then tried to reconstruct verbatim the serial order of the words he had studied. This reconstructed order was then given to a second subject to study and recall, and his reconstructed order was then studied by a third subject and recalled, and so on. The passage was successively filtered through 16 subjects. As it went from subject to subject, its recall (reconstructed order) was distorted more and more from the original jumble in the direction of sensible English sentences. One of Coleman's original passages and the sixteenth reconstruction of it is given in Table 6.3. The amount of change is dramatic, especially considering that each sub-

TABLE 6.3. Change in recall of a passage as it moves through a chain of learners. (Adapted from Coleman, 1962.)

Original passage studied by first subject:	"about was good-looking way and treating made of that a him the quiet youngster nice he manners a them girls wild go with . . ."
Reproduction of sixteenth subject in the chain:	"he was a youngster nice quiet with manners good-looking and a way of treating them that made the girls go wild about him . . ."

ject was trying to reproduce verbatim the exact order of words he had studied. The change illustrates vividly the powerful effect of prior verbal habits in distorting recall of conflicting associations.

Interference with Meaningful Text

The laboratory studies of interference reviewed above have used "meaningless" materials, either nonsense syllables or random unrelated words. However, recent research provides extensive evidence that similar interference processes operate in the learning and forgetting of meaningful text materials, both at the level of single sentences and at the level of interrelated sets of sentences (text paragraphs). Although some doubts about interference processes had been raised by a few early nonanalytic experiments, recent positive demonstrations show retroactive interference or retroactive facilitation depending in a very lawful manner upon the exact arrangement of materials and the retention measure used for assessing losses (see J. R. Anderson & Bower, 1973; R. C. Anderson & Myrow, 1971; R. C. Anderson & Carter, 1972; Crouse, 1971; Myrow & Anderson, 1972).

We may illustrate the issues with recall of simple active declarative sentences of the form subject-verb-object (e.g., "The mechanic repaired the refrigerator"). We may view this proposition in either of two ways: the sentence establishes in memory a serial chain of associations between the successive words, or it establishes labeled functional connections between groups of semantic concepts which are aroused by these specific words. A variety of considerations suggest that the latter view is more nearly correct and more fruitful. We may now treat the semantic concepts corresponding to the subject, transitive verb, and object (call them S, V, O) as though they were terms in a "triplet" association-learning task, except that there are a tremendous number of syntactic constraints and semantic selectional restrictions which forbid certain word combinations (e.g., "Night the bounced running the" is literal nonsense).

Treating the concepts as terms in an associative triplet, we are then led to expect interference at this level if a given concept co-occurs with different concepts in new predications. Thus, following learning of S_1-V_1O_1, interpolation of S_1-V_2O_2 or S_2-V_1O_1 will lead to an associative loss between S_1 and V_1O_1, when the cueing is done in either direction. This loss would be assessed relative to a control interpolation symbolized as S_2-V_2O_2. An experiment by G. H. Bower (1978) demonstrated clear negative transfer and retroactive interference by subjects learning interpolated sentences bearing an *A-D*, *C-D*, or *A-Br* relation to the original sentences learned (see Table 6.1 for transfer conditions). The results were ordered exactly as predicted by paired-associate transfer results.

Further, interference evidently occurs at the level of conceptual learning. This can be demonstrated by using synonymous paraphrases, say of the V_1O_1 verbal construction. Suppose S_1-V_1O_1 is the sentence "The sheriff aroused the slumbering patient"; letting $P(V_1O_1)$ denote synonymous paraphrase of the verb phrase, then an S_2-$P(V_1O_1)$ sentence might be "The nurse awakened the sleeping sick person." It has been found that interpolation of such paraphrase constructions produces nearly as much retroactive interference as does use of the verbatim V_1O_1 paired with a new subject-noun, S_2 (see R. C. Anderson & Carter, 1972). Apparently, the cue V_1O_1 contacts a similar trace in memory as does its paraphrase $P(V_1O_1)$, and the S_2 associated with $P(V_1O_1)$ competes with recall of the S_1 associated earlier with the same predicate by use of the words V_1O_1.

Of course, this paraphrase effect could

be used to good advantage if one wished to facilitate conceptual (meaning) associations. Thus, a two-list experiment in which S_1-V_1O_1 in List 1 is followed by S_1-$P(V_1O_1)$ in List 2 will result in enhanced recall of the correct gist (meaning) to the S_1 cue of List 1, but probably with some loss in verbatim recall of V_1O_1. For the same reason, a paraphrased sentence $P(S_1)$-$P(V_1O_1)$ (e.g., "The policeman awakened the sleeping sick person") will be readily learned with high positive transfer following learning of S_1-V_1O_1. It is important to remain clear on this distinction between verbatim and gist recall, since it is possible to facilitate associations between general concepts (gist) at the same time one is interfering with verbatim recall.

A particularly striking example of the separation of these two levels was provided in an experiment by S. A. Bobrow (1970). He showed that if the subject and object nouns of a sentence were repeated as a pair in a second sentence, the association between them could be enhanced or not depending on whether the meanings of the nouns were maintained across the two sentence contexts. To illustrate, subjects learn to associate pairs of nouns like *pitcher* and *jam*, and do so using linking sentences as mediators. Suppose a sentence in the initial study list were "The milk *pitcher* was splattered with sweet *jam*." A sentence in the second list which preserved a similar conceptual meaning for the nouns would be "The lemonade *pitcher* was sticky with strawberry *jam*," whereas a sentence which totally altered the conceptual meaning would be "The baseball *pitcher* got caught in a traffic *jam*." The retention test, given at the end of the second study list, involved presentation of the subject noun (*pitcher*) for recall of the object noun (*jam*). As our intuition suggests, the subject-to-object association was greatly enhanced by interpolation of an identical or conceptually similar predication. But there was no accumulative learning when the meanings of the words were changed; performance was similar to what would have occurred if the *pitcher-jam* pair had been presented only once.

Although the illustrations above are for recall of single, isolated sentences, interference effects have been shown also for paragraphs, stories, "science lessons," and the like. In such experiments, one must attend carefully to what are the atomic assertions that relate concepts in the initial text, and how the specific predications about these concepts are altered in the interpolated learning. For example, Crouse (1971) and G. H. Bower (1974) have used short biographies of fictional persons as experimental passages. These comprise essentially a listing of life-history facts about the person. An interpolated passage which will produce little interference with the biography might concern, say, an art exhibit, whereas a passage causing maximal interference would be a second biography which systematically changes some of the facts contained in the first biography (such as names, dates, places, occupations, and so on). Bower found that specific facts which remained the same in the two biographies were facilitated in recall of the first biography whereas specifics which were changed (e.g., father's occupation) were forgotten, even though the subject was likely to remember to say something about the right general class of facts. That is, the subject might remember to say something about the father's occupation but would get the specific details wrong. This suggests, as mentioned earlier, that by appropriate interpolation one can selectively facilitate recall of the conceptual "macrostructure" of a passage while at the same time interfering with memory for the specific "microstructure" of the material. A later experiment by Thorndyke and Hayes-Roth (1979) is especially clear in showing these two effects—the learning of the macrostructure alongside interference with the microdetails of the passage.

SUMMARY COMMENTS

This brief tour through interference theory will suffice to indicate its major features. The main shifts that have occurred in it have been acceptance of the notions of unlearning and of response-set suppression, a new emphasis on proactive interference, and identification of a potent source of proactive effects in those prior language habits that conflict with the temporary verbal associations set up in a laboratory experiment. Increasingly, research is being directed at understanding interference and the forgetting of meaningful sentences and larger bodies of text. There have also been changes in the experimental techniques employed. For example, Barnes and Underwood's modified recall procedure mentioned earlier is now widely used because of the additional information it yields on what the person remembers. Pair recognition is used to assess associative learning, whereas free recall is often conceived as an index of pure "response availability." This is an active field of research and no brief discussion can do justice to the range of variables that have been investigated in relation to forgetting. For more comprehensive reviews, see McGeoch and Irion (1952), J. F. Hall (1971), Postman (1971), and Spear (1978).

Current studies of human learning and forgetting have continued in a highly analytical phase, with attention on progressively finer analysis of smaller aspects. As happens during analytical phases in other specialties, synthesis of the knowledge into a broader conception of the phenomena has been shunted aside. As a result, the possible uses of our scientific knowledge for solving practical problems have been only cursorily explored, and then in an often stumbling fashion. To mention just one major applied problem: educators or anyone engaged in training personnel would surely like to know how best to teach students something so that they will retain it for a long time. The laboratory work relating retention to the conditions under which training has occurred is clearly relevant; but it is often so far removed from the kind of task, background, and other variables that make up the applied situation that some ingenious extrapolation is required before the principles can be put to use. Writings by Gagné (1970), Staats (1968), and collections of papers edited by Hilgard (1964a) and by DeCecco (1967) represent a few attempts at reasonable extrapolations. The area of instructional psychology aims to improve instructional practices in schools by use of learning principles, task analysis, and skill training.

Characteristics of the Verbal Learning Tradition

Having reviewed research around selected topics in the verbal learning tradition, what may we say about the characteristics and orienting attitudes of the area's practicing researchers?

First, this functionalist tradition was committed to a firm environmentalism, believing that individual differences have largely arisen from differences in acquired skills and habits. They believed strongly in historical causation of current behavior; the cause of the person's present responses are to be found in her past training and how it is being transferred to the present situation.

Second, "mind" is considered as a collective name for a set of *dispositions* to behave in particular ways in particular circumstances. To describe a person's mind as bright or dull, retentive or forgetful, quick or slow is not to refer to some inner entity that causes individuals to act in certain ways; rather, it is to refer to their abilities and tendencies to act in these characteristic ways. This position is practically the same as the behaviorist's program.

Third, the verbal learning theorist prefers scientific concepts that are intersubjectively countable or measurable. Thus, the "meaningfulness" of a word will be identified with the average number of associates it evokes in thirty seconds, although this bypasses traditional definitions of meaning in terms of reference, use, or defining properties. Verbal learning theorists have a bias against mentalistic constructs such as imagery and nonverbal thinking. Insofar as possible, they assume that perceptual events are categorized, coded, and stored according to the verbal labels elicited by the events. Verbal learning theorists also oppose vague, heuristic constructs such as organization, structure, insight, and the gestalt properties of stimulus sets. Whenever possible, they translate the fuzzy and the mystical into simpler terminology of habit repertoires or stimulus patterns. Discontinuities in learning are believed to hide underlying continuities in habit acquisition.

Fourth, as noted, the verbal learning tradition has been firmly committed throughout to associationism, with a prolonged romance in midlife with Hull's S-R theory, since left behind in detail if not in spirit. Verbal learning theorists have tried to explain practically all phenomena uncovered in the verbal learning laboratory in terms of particular associative networks established between stimuli and responses, and the operation of simple retrieval rules such as cue similarity and response competition. The general tenets of associationism were criticized in Chapter 1 and shall be again in Chapter 13.

Fifth, the verbal learning tradition has been carried forward by a small group of psychologists, their students, and their students' students. Influential early functionalists (after Ebbinghaus) were Carr, Dewey, and Woodworth; they influenced McGeogh and Irion, who influenced Melton, Bilodeau, Underwood, Cofer, Osgood, and Postman, who influenced Keppel, Schulz, Spear, Martin, and so on. The research fervor for particular issues has been passed between generations, with many students of these scientists going on to productive careers in their own right. The dedication and productive energy of this group of researchers has earned them the gratitude of their scientific colleagues.

Modifications in Associationism

It is fair to say that associationism in the hands of modern verbal learning theorists is a different animal from the older British philosophical associationism, although it has evolved out of that philosophical tradition in response to critiques and empirical findings. Several of the criticisms are cited in Chapter 1 (under "Rationalism") and we will review a few of them here. First, while classical associationist theory said little about perceptual organization of sensory elements, it is now clear that the organization and "belongingness" of the sensory material greatly affects what the subject learns. Thus, when instructed to listen to a sequence of word pairs, *A-B, C-D, E-F* the hearer will segment the stream into pairs, and an element will become associated to the other members of its pair but not at all to elements of preceding or following pairs despite their objective temporal contiguity. Modern associationists accept such results but try further to show that the segmentation or grouping operation itself may be viewed theoretically as an attentional strategy or learnable higher-level response, which will show negative transfer, and so on.

A second point against classical associationism is that it did not recognize the many different *types* of associations that encode different types of relationships between two ideas. Labeling associations according to their type would permit efficient, relation-directed searches of memory and allow the direct retrieval of answers to questions like "What concept has relation R to concept X?" (e.g., "What's the superordinate category for *canary*?"). The retrieval of answers to such questions can be so fast that it is difficult to believe that individuals are somehow looking through long lists of associations in their memories trying to find an element on a list of superordinate names that is also on the list of associates to their concept of *canary*. Direct, relation-guided retrieval would seem more consonant with fast retrieval.

Related to the above is another complaint against classical associationism, which is that the classical theory does a poor job of explaining why the association retrieved can depend so much on the context in which a stimulus occurs. Thus, a red light means "stop" as a traffic signal but means "go through here" as a fire-exit sign. But this problem can be dealt with by assuming that retrieval cues always act within a patterned complex, that different goals or environments lead to different stimulus components in short-term memory (goal A or B or . . .), and that a specific stimulus (X) arouses different associations in the pattern $A + X$ than in the pattern $B + X$.

A further point not anticipated in classical associationism is that many of the associations observed (such as *RZL-CAT*) are not in fact direct ones but rather are mediated through a chain of more elementary associations. Thus, the temporal contiguity of *RZL* and *CAT* in the learner's experience has not resulted in her setting up an independent association, but rather has initiated a memory search for familiar associations which would solve the problem posed for her—namely, to unify arbitrary pairs of units. Of course, the notion of chains of mediating associations is within the spirit of modern associationism.

A third problem with classical associationism is that only simple ideas were assumed to enter into associations, and associations were supposed to link only simple ideas, not chunks of ideas. Hull maintained a similar restriction. This restriction says that only "horizontal" links are permitted in theory (see Wickelgren, 1979b, for a discussion). This associative structure and two alternative ones are shown in Figure 6.15 for four interassociated ideas A, B, C, D. Panel A shows the classical hypothesis, that only direct element-to-element connections are permitted. (For this diagram, interpret the links as two-way associations.)

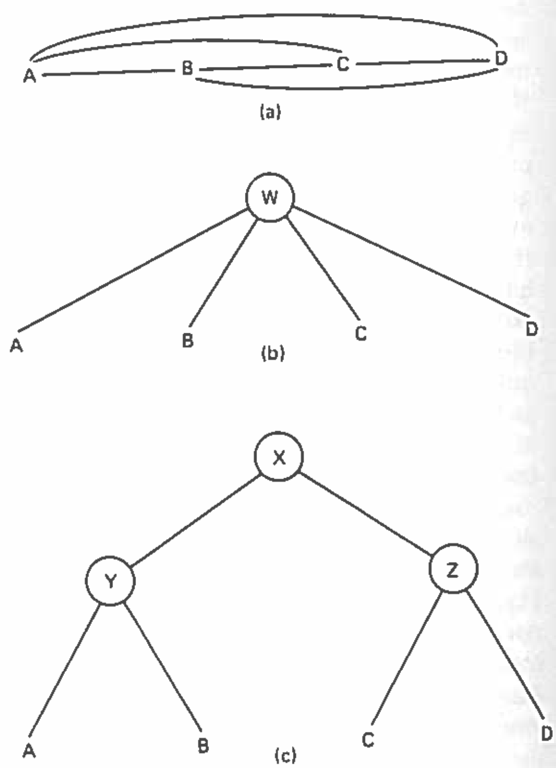

Figure 6.15. Illustration of three types of associative structures for interrelating four items or properties, denoted A, B, C, D. (A) shows only horizontal associations. (B) permits vertical associations to the chunk node, W. (C) shows vertical associations outlining a hierarchy of two groups or chunks (Y, Z) that are grouped into the higher-level chunk, X.

Panel B shows a structure acquired by creating an abstract chunk or group node, *W*, into which each of the lower elements *A, B, C, D* has two-way links. Here, the ideas *A, B, C, D* are not directly linked to each other, but "communicate" with each other only through the chunk-node, *W*. The node *W* is abstract in the sense that it represents nothing except the co-occurrence of the four component ideas as a pattern or group; node *W* can also be thought of within an associative network as a switching terminal for shunting excitation arriving at *W* from a retrieval cue, say, cue *A*, out through the several pathways to *B, C, D*, thus to activate those ideas and bring them into consciousness. This terminology makes manifest that an associative network is basically a set of communication pathways, with nodes defined according to which other nodes (inputs) activate them and the nodes (outputs) to which they transmit activation.

Experiments by Ross & Bower (in press) and Arnold (1976) have supported the predictions of the vertical model in Panel B in preference to the horizontal model in Panel A. In the former study, adults studied many clusters of four or five slightly related words, then were tested for recall of each cluster by cueing with one or two words. The frequencies of the several recall patterns given one and two cues were fit quantitatively far better by the vertical model. Arnold (1976) came to a similar conclusion examining one-cue and two-cue recall of studied word-triplets as well as recognition of pairs and triplets. Thus, the evidence suggests the vertical model is preferable to the horizontal one. Anderson and Bower (1973), Estes (1972), and Wickelgren (1979) among others have also suggested the vertical association model.

Panel C in Figure 6.15 illustrates that chunk nodes (*Y, Z*) themselves can be grouped into a higher-order node (*X*), giving one the capability of hierarchically organizing base elements into groups. Clearly we need a knowledge representation that permits any segmentation and grouping of elements, and also one that allows us to represent recursive grouping. This is required, for example, if the theory is to represent the person's memory for series of groups of elements. Such a representation is assumed in theories such as Lesgold and Bower's (1970), Estes's (1972), or Johnson's (1970) accounts of memory for chunked serial lists. Anderson and Bower (1973) also used a hierarchy of segmented idea units to represent complex sentences in memory. Thus, for example, in a simple declarative sentence, the noun phrase ("The old man") might correspond to node *Y* in Figure 6.15c, and the verb phrase ("petted the dog") might correspond to node *Z*; then, the top-level node *X* would be the internal code in memory permitting access to the conceptual structure established by the subject hearing and understanding the assertion "The old man petted the dog." These higher-level nodes like *X, Y, Z* can themselves enter into further associations. Thus the expanded associative theory envisions the growth of arbitrarily complex concepts encoded as associative configurations of elements, groups of elements, and groups of groups. These network formalisms underlie much of the developments in neo-associative theories of knowledge, concept utilization, and positional learning (Anderson & Bower, 1973; Kintsch, 1974; Norman & Rumelhart, 1975). These developments will be reviewed in Chapter 13.

Criticisms of the Verbal Learning Tradition

Critics of the verbal learning tradition have pointed to several shortcomings. First, it has used associationistic concepts of memory storage to understand its results, but has typically failed to specify the "executive monitor" which uses that memory base for answering questions or solving problems. As one illustration, verbal learning psychologists have proposed the idea that subjects edit their recall in a multi-list experiment when they are trying to recall a specific target list; as each item is retrieved, it is checked for a target list tag and is suppressed if it comes from the wrong list. But in a strictly associative theory, *what* carries out this editing function? How do we represent the program or routine that has been installed in short-term memory to guide this "generate-candidates-then-test" strategy for recall? How are those programs acquired? Typically, traditional verbal learning psychologists have bypassed such questions. Recent computer simulation theories such as J. R. Anderson's ACT model (1976) deal explicitly with these executive processes which use the associative memory.

A second common complaint is that the studies in the verbal learning tradition have nothing much to say about language learning. But verbal learning grew up during the "conditioning era" of psychology, was always intended to study basic associative learning in simplified arrangements, and was never aimed at bringing realistic school tasks directly into the laboratory. Many school tasks clearly do have large rote-learning components, and for these it appears that the results of the laboratory analogs apply as expected (for example, interference among similar biographies).

A third complaint of the critics is that verbal learning research is "crassly empirical," that it generates bushels of detailed data without revealing powerful, general principles or theories. To that criticism, the verbal learner would counter that science is first of all analysis and description of any phenomenon in all its myriad facets, and that "grand theories" are simply wasteful "grand delusions" unless one first has a secure empirical footing for theorizing. One should first try to figure out most of the variables that could influence some experimental phenomenon in order to make informed guesses about theory.

Verbal learners have proposed and researched a large number of theoretical hypotheses, although these tend to be local, tailored to a restricted domain, and have limited boundary conditions. But in terms of sheer hypotheses alive and well, the verbal learning psychologists can hold their own with others working on theories of animal learning. As global theories have lost their allure, as miniature hypotheses are being developed increasingly for specific learning tasks, the eclectic stance of functionalism has an increasing appeal for contemporary experimental psychologists.

SUPPLEMENTARY READINGS

The following books in the functionalist tradition may be recommended:

Bilodeau, E. A. (1966). *Acquisition of skill.*

Hall, J. F. (1971). *Verbal learning and retention.*

Kausler, D. H. (1974). *Psychology of verbal learning and memory.*

Kling, J. W., & Riggs, L. A. (1971). *Experimental psychology.*

McGeoch, J. A., & Irion, A. L. (1952). *The psychology of human learning.*

Melton, A. W., ed. (1964). *Categories of human learning.*

Osgood, C. E. (1953). *Method and theory in experimental psychology.*

Robinson, E. S. (1932a). *Association theory today.*

Spear, N. E. (1978). *The processing of memories: Forgetting and retention.*

Underwood, B. J. (1966). *Experimental psychology.* 2nd ed.

Woodworth, R. S. (1958). *Dynamics of behavior.*

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology.*