3

Optimizing Long-Term Retention and Transfer

This chapter considers training conditions that do or do not facilitate posttraining performance. We focus on two aspects of posttraining performance: its durability (long-term retention), that is, the extent to which a training program yields a level of learning that supports performance after long periods of disuse; and its flexibility (transfer), that is, the extent to which a training program prepares a learner to perform under real-world conditions that may differ from those present during training. Our primary concerns are the training of adults rather than children and the training of cognitive-motor procedural skills, such as programming a computer, repairing a mechanical or electronic device, hitting a serve in tennis, parachuting out of an airplane, or receiving and transmitting Morse Code, rather than on classroom learning.

Focusing on the training of adults and on procedural tasks helps to limit the scope of this chapter, but covering all aspects of training so defined in one chapter is still prohibitive; a broader view is provided in the recent book by Farr (1987). Among the issues we do not address are individual differences among learners, instructor variables, such as motivating trainees and improving instructor-trainee rapport, and technological innovations, such as computer-based instruction and the use of simulators; those issues are addressed in recent reviews by Montague (1988) and Walberg (1990), as well as in Farr (1987). We also do not focus on how the individual components of complex tasks should be integrated. (Chapter 4 takes a broader view of complex cognitive tasks and includes a discussion of some of the ways computers can assist training.) Even with our mission defined more narrowly, the relevant literature deriving from basic and applied research in education, psychology, cognitive science, physical education, and sports psychology is enormous: thus, the references we cite should be viewed as representative, not exhaustive.

Procedural knowledge is now commonly distinguished from declarative knowledge (Winograd, 1975). Declarative knowledge is knowledge of facts or static information (e.g., in what year did Babe Ruth hit 60 home runs?), and it, in turn, is typically subdivided into episodic and semantic knowledge: episodic knowledge is context dependent, such as knowing what you ate for lunch today; semantic knowledge is independent of context, such as knowing what Europeans typically eat for breakfast (see e.g., Tulving, 1985).

Procedural knowledge is knowing how to execute the procedures necessary to perform a given task. Procedural knowledge underlies cognitive and motor skills (many of them automated), such as how to change a flat tire on a car, use a typewriter by touch, operate a computer, disassemble and reassemble a rifle, ride a bicycle, or play a game. Skills are acquired mainly by doing or practice and are not learned quickly. Retention of skills, or the lack thereof, is typically measured by the extent to which they can be performed, rather than by the extent to which they can be "recalled" per se. In fact, at high levels of skill, in which many of the procedural components of a skill become automatic, people become unable to describe in any detail what procedures they are carrying out in what order. A person may have to resort to consciously observing his or her own behavior, for example, to tell a friend how to ski or operate a standard-transmission automobile.

It seems obvious that the major goal of any training program is to prepare trainees to perform effectively on a posttraining task in a real-world setting; achieving that goal, however, is complicated by several factors. First, what is observed by those responsible for training programs is, typically, the performances of trainees during the training process itself. Such performance is a highly imperfect index of the kind of learning, comprehension, or understanding that will sustain performance of the skill or knowledge over periods of disuse. Someone who meets high standards of performance at the end of training may fail to perform adequately some months later. Acquisition of a given skill during training also does not provide evidence that the learner will be able to perform in contexts that differ from the training context or on altered versions of the training task that may arise in real-world settings. The term "context" includes the task, practice conditions, and cognitive processing used by a trainee.

The crux of the problem is that learning and performance are not the same. As we indicate at several points in this chapter, procedures that enhance performance during training may or may not enhance long-term retention and transfer to altered contexts; conversely, procedures that introduce difficulties for the learner and impair performance during training may foster durable and flexible posttraining skills (for some examples, see Schmidt and Bjork, 1992). In short, the goal is to have training programs that optimize learning—some relatively permanent change in the capacity for responding—but what is observed during training is performance localized in a given place and time. At a later time, in another place, the learner may perform quite differently and that performance is often at an inadequate level. The performance observed during training may be mediated by rote memory or cues specific to the training procedure rather than being indicative of any substantial learning or understanding.

This problem is aggravated in training settings in which those who are responsible for training do not see the posttraining performance of the individuals they have trained. In such a setting, the instructor's judgment as to the efficacy of different training procedures may be governed entirely by the tacit assumption that what yields high performance during training will yield high retention and transfer after training. In any organization in which the people responsible for the maintenance of critical skills and knowledge (refresher training, retraining, and so forth) are not the same people who are responsible for initial training—the military is such an organization—this inferential problem is going to be particularly troublesome.

LONG-TERM RETENTION

When one assesses posttraining performance on some task, the time interval from the end of training to the performance "test" can be varied, the task can be the same or an altered version of the training task, and the situational context can be similar to or different from the training context. Thus, someone trained to repair a certain type of pump in a nuclear power plant might attempt the first such actual repair many weeks or months after being trained, the pump may differ in certain respects from those encountered in training, and the repair may need to be executed under conditions of heat or other pressure that was not present during training. It is common to speak of retention when performance on the actual training task is assessed under posttraining conditions that are essentially the same as the training conditions. The term transfer is used when the posttraining task or setting differs from the training task or setting. For convenience, and to be compatible with the literature, we tend to follow that usage, but it is important to emphasize that retention so defined is a special case of transfer. That is, since the posttraining context will never match exactly the training context—if for no other reason than that the physical, emotional, and mental state of the learner will not be exactly the same—a test of retention can be viewed as a test of the transfer of training to contexts that appear to match the training context.

In attempting to make our review compatible with the literature, in which the learning-performance distinction is often blurred or forgotten by researchers, we often need to speak of the level of "learning" achieved during training when level of performance would be the more correct expression. Terms such as "original learning" and "overlearning" are too common for us to avoid. In the next section particularly, when we discuss retention and transfer as a function of the level of original learning, we have tried to restrict our coverage to research situations in which it can be generally assumed that the performance levels measured as evidence of differing levels of learning do, indeed, denote just that. In later sections we deal with training situations in which performance during training is a poor measure of the level of learning achieved.

Original Learning

There is considerable agreement that the long-term retention of a task can be improved by increasing the level of original learning or mastery (e.g., Annett, 1979; Farr, 1987; Gardlin and Sitterley, 1972; Hagman and Rose, 1983; Hurlock and Montague, 1982; Naylor and Briggs, 1961; Schendel et al., 1978; Prophet, 1976). Indeed, the level of original learning for a task is the best single predictor of long-term retention for any given retention interval. Thus, any variable that can help trainees achieve a higher level of original learning or mastery of a task is capable of enhancing its retention (Hurlock and Montague, 1982).

Most often, the training variable manipulated is the amount of practice on a task. Typically, this manipulation is accomplished by making the criterion of mastery more difficult to achieve so that more practice is needed to achieve the criterion. The additional practice needed to achieve the more difficult criterion produces a higher level of original learning, which enhances retention. For example, suppose a basketball coach is training young, novice players to shoot free throws (foul shots). He or she decides that all of the players should be able to make 25 out of 50 shots attempted by the end of 8 weeks of training, which is a reasonable criterion of mastery. But the coach could also make the criterion of mastery more difficult to achieve, requiring that the players be able to make 35 out of 50 shots, or, alternatively, that the players be able to make 25 out of 50 shots under more difficult conditions (after wind sprints, with simulated crowd noise, at alternate baskets, and so forth).

In either case, more practice would be needed to achieve the more difficult criterion, but the additional practice would produce a higher level of original learning, which, in turn, should lead to greater retention.

Setting the Criterion of Mastery

Determining the appropriate level of original learning of a task that trainees must acquire in order to ensure the desired level of retention is not easy. How the level is selected—that is, how it is defined and assessed in terms of when the criterion of mastery is satisfactorily achieved—is quite arbitrary. Typically, the level is measured in terms of trials or time taken or number of errors committed until the criterion of mastery has been achieved. In most of the studies the committee reviewed, the criterion has been set either at a minimal mastery level, such as the first errorless performance trial, or at somewhat more than minimal level: for example, the three successive errorless trials required by Rigg and Gray (1981) in their research on U.S. Army enlisted personnel learning a procedural task. More recently, Jones (1985) suggested another way to determine when the level of original learning is satisfactorily achieved. He recommended that in addition to selecting some arbitrary criterion of mastery, the acquisition curve be used to determine when performance has stabilized at or above the criterion level. When the slope of the acquisition curve has begun to level off above the criterion level, learning would be considered complete.

Automaticity

The level of original learning can also be assessed in terms of the degree of automaticity of performance by using a dual-task paradigm, in which a secondary task is given to trainees to sample their spare cognitive capacity while they are learning the primary task (Shiffrin and Schneider, 1977; Schneider and Shiffrin, 1977; Schneider et al., 1984). An acceptable degree of automaticity and, hence, level of original learning, is the point at which neither the primary nor secondary task causes a performance decrement on the other. Although the dual-task paradigm is an acceptable method of assessing automaticity, it is not without its problems (see, e.g., Fendrich et al., 1988; Jonides et al., 1985). Theoretically, skills that require only a minimum of attention and cognitive capacity to perform are either completely or partly automatic, whereas skills that require cognitive resources and effort involve controlled processes. Schneider et al. (1984) define an automatic process as one that does not make use of general cognitive resources. In other words, capacity reductions do not influence automatic processing. Moreover, an automatic process is not subject to conscious control and, thus, can be executed in response to relevant external stimuli to which little attention is paid.

Whether or not a skill is classified as automatic or controlled depends largely on the level of original learning. Many skills require controlled processes early in learning, but the processes become automatic with extensive practice and especially so if that practice contains a high degree of consistency. Practice consistency means that a trainee makes the same response each time a certain stimulus or class of stimuli is presented. The assumed explanation for this phenomena is that retention of a skill depends heavily on the extent to which a skill is automatic: that is, it can be performed without conscious awareness. The more automatic the skill, the greater the chance that the skill will be retained over nonuse periods without refresher training or rehearsal. It is important to note, however, that the automated parts of a skill that are acquired through practice (e.g., speed in soldering a joint) are expected to deteriorate during nonuse periods, while the automated parts of a skill that are less dependent on practice (e.g., encoding of temporal or spatial information) are not. Thus, for designing a skill maintenance program, more emphasis should be placed on the automated parts of the skill that are acquired through practice.

Overlearning

Regardless of the criterion of mastery selected for original learning, one way to enhance retention is to provide supplementary practice on a task after the criterion is achieved. In the previous basketball example, for instance, the coach could have the players continue to practice freethrow shooting even though they have achieved the criterion of making 35 out of 50 shots. This method may be interpreted as postmastery learning and is usually referred to in the literature as overlearning. Level of overlearning is usually expressed as simply the number of practice trials that trainees perform after the criterion of mastery has been achieved, or it is expressed in percentage terms—50 percent overlearning, for example, means that trainees receive half again the number of trials that they took to achieve the mastery criterion. The arbitrary nature of mastery and overlearning criteria can make it difficult to do certain comparisons across studies. A trial that is part of mastery for one study can be part of overlearning for another study. It depends on how one defines *when* original learning is complete, how one *quantifies* the level of original learning or mastery, and how one *defines* the level of overlearning.

Those complications notwithstanding, it is clear that retention is better for overlearned tasks (e.g., Loftus, 1985; Schendel and Hagman, 1982; Slamecka and McElree, 1983). For enhancing retention, when to introduce the supplementary trials does not appear to be a critical factor; the level of overlearning is far more important than the time at which the supplementary trials are introduced (Schendel and Hagman, 1982). The literature also reveals, however, that providing overlearning trials reaches a point of diminishing returns (e.g., Bell, 1950; McGeoch and Irion, 1952; Melnick, 1971). In other words, increasing the number of overlearning trials may not produce proportionate increases in retention. Thus, although 100 percent overlearning may result in better retention than 50 percent, the additional gain that occurs may not be worth the additional time and practice.

Task Cohesion and Organization

In the literature on task retention, a given task is often classified as a discrete motor task, a continuous motor task, a procedural task, or a verbal task. Discrete tasks are characterized by a clearly defined beginning and end, such as responding to a signal by pressing a lever or saying a word. Continuous motor tasks require responding to information presented continuously—such as driving a car, which is a familiar example of a tracking task. Procedural tasks consist of a particular sequence of operations executed in the same way each time that the task is performed, such as disassembling and reassembling a rifle. Verbal tasks involve materials ranging from letters to nonsense syllables to words to meaningful prose, such as doing a crossword puzzle.

Considerable attention has been devoted in training to procedural tasks, probably because they are so easily forgotten and are common in almost all work situations. Such tasks may vary on several dimensions, such as the number of steps they contain, the degree to which performing one step cues another, the freedom to vary from a fixed sequence, the extent of planning required to execute the task, and the number of decision points. Shields et al. (1979) found that the rate of forgetting for a procedural task is predominantly a function of the number of steps needed to perform the task. Moreover, they found that what tended to be forgotten most were the steps not cued by equipment or by the preceding steps.

Consistent with such an argument, there is consensus in the literature that continuous motor tasks are better remembered than discrete or procedural tasks mainly because the former have a higher degree of inherent organization. Prophet (1976) proposed that the poor retention exhibited for certain procedural tasks—such as instrument flying—was due primarily to a low degree of internal organization or cohesiveness. Thus, regardless of the type of task, it appears that it is the degree of organization or cohesiveness of the task that is a principal determinant of the level of original learning that is achieved and the amount that is retained in the long term (e.g., Hagman and Rose, 1983; Hurlock and Montague, 1982; Prophet, 1976; Schendel et al., 1978).

Enhancing Retention

Distribution of Practice

We have focused thus far on amount of practice during training, and have assumed that the level of performance achieved during training is a reasonable index of the level of learning achieved. For a fixed amount of practice, however, learning (as measured by a later retention test) depends on the temporal distribution of practice, and the nature of that dependency illustrates that performance during training is an unreliable indicator of learning. In general, massing of practice on some component of the to-be-learned task produces better performance in the short term (e.g., during training) but much poorer performance in the long term than does spacing of practice. In some cases, massed practice yields long-term recall performance less than one-half the level that results from spaced practice, and two massed practices are often not appreciably better than a single study trial (see, e.g., Glenberg, 1979; Glenberg and Lehmann, 1980; Melton, 1970; and Rothkopf and Coke, 1966).

The so-called spacing effect—that practice sessions spaced in time are superior to massed practices in terms of long-term retention—is one of the most reliable phenomena in human experimental psychology. The effect is robust and appears to hold for verbal materials of all types and for motor skills (for reviews, see Crowder, 1976; Dempster, 1990; Lee and Genovese, 1988). A recent indication of how durable the advantages of spacing may be across truly long posttraining intervals was reported by Bahrick and Phelps (1987). They tested subjects' recall of English-Spanish word pairs 8 years after the original training phase. During the training phase, successive practice sessions were separated by 30 days, 1 day, or 0 days. The level of retention was highest for the 30-day spacing of study sessions, next highest for the 1-day spacing, and lowest for the 0-day spacing, with performance for those in the 30-day condition more than twice that for those in the 0-day condition.

Given the benefits of spaced practice and the fact that those benefits have been known to researchers since the beginning of controlled research on human memory (Ebbinghaus, 1913), one would expect that spaced repetition would be a major component of modern programs of training and instruction. The fact that this not seem to be the case is something of a puzzle (see Bjork, 1979; Dempster, 1990). Part of the solution to that puzzle, of course, may be a point we have already stressed: during the training process itself, spaced practice may appear inferior to massed practice.

Another factor in the apparent neglect of scientific findings on distribution of practice by those responsible for the design of training programs is time pressure: massed sessions take less total time than do spaced sessions. A study by Baddeley and Longman (1978), carried out for the British Post Office, illustrates the point. Given a new sorting system, which required postal workers to enter postcodes into a sorting machine using a standard typewriter keyboard, a large number of postal workers needed to be taught to type in a relatively short period of time. Baddeley and Longman examined four different training schedules, ranging from 1 hour of practice per day (spaced) to 4 hours of practice per day (massed). In terms of the learning curve—a plot of mean keystrokes per minute as a function of hours of practice—spaced practice was far more efficient than massed practice. To reach any given level of performance, however, it took the 1-hour-per-day group many more days than it took the 4-hours-per-day group, and the authors report that the former group was the least satisfied because the members felt they were falling behind the groups that were getting more practice per day. Thus, spaced practice produced much more efficient learning as a function of time on task, but took more days, which could certainly be a negative factor from a management standpoint.

Fostering Understanding

Just as the organization or cohesiveness of the components of a task makes it easier to learn and remember, so too does the organizing influence of understanding (Horton and Mills, 1984; Wertheim, 1985; Wetzel et al., 1983). In a story, independent or vaguely related occurrences are similar to steps of a procedural task that are not logically arranged and, hence, do not signal each other. When relevant organizing information is provided before reading a fragmented story, this information supplies a coherent structure within which to interpret more effectively the exact meaning of the story (Owens et al., 1979). Moreover, when this structure is also compatible with a trainee's general knowledge of the world, recall is enhanced (Morris et al., 1979).

There is considerable evidence suggesting that long-term retention of procedural tasks that are based on complex rules or principles can be enhanced by augmenting instruction with explanations or information designed to increase a learner's understanding of the to-be-learned tasks (Gentner, 1980, 1982; Smith and Goodman, 1984; Tourangeau and Sternberg, 1982; Kieras, 1981; Sturges et al., 1981). Although researchers differ to some extent on how they define explanations, it seems useful to categorize them as linear, structural, and functional (Stevens and Steinberg, 1981; Smith and Goodman, 1984). Linear explanations tell a trainee what to do—that is, what steps to follow and in what order. Structural explanations clarify how or why different task components belong together. Functional explanations inform the trainee about the cause-and-effect relationships among task components. In general, linear and structural explanations are used for static tasks, such as assembling a piece of equipment; functional explanations are used for dynamic tasks, such as operating a piece of equipment. In an examination of some of the literature dealing with the long-term retention of conceptual information and procedures inherent in expository prose as a function of structural explanations, Konoske and Ellis (1985:13) conclude that effective structural explanations "should include spatial and component-part information . . . as well as . . . goal statements. In addition, structural information should be communicated using text, schematics, graphs and illustrations, whenever possible."

In another study, Kieras and Boviar (1984) provided subjects in an experimental group with a "mental model" functional explanation of a new device that they were required to learn to operate: a mental model is an internal conceptual structure that corresponds to some aspect of the world (see, e.g., Gentner and Stevens, 1983). The subjects in the control group were not provided with any explanation; they had to learn the procedures solely by memorizing them. The researchers found that the functional model was more effective for enhancing retention one week after original learning than was learning by rote. They explained the superiority of the functional model by claiming that it was more pertinent to the operation of the device and could be used to cue operational procedures that might not otherwise have been retained. This finding and explanation supports Farr's (1987) position that the long-term retention of procedural knowledge and skills can be greatly enhanced if trainees understand why tasks must be performed in a particular order and way; the relationship of the parts to the whole task; and how new task information is related to what is already known. Farr (1987:78) claims:

Understanding enables the trainee to (a) furnish himself with cues to help retrieval; (b) recognize the relationship of externally provided or system-provided cues to the sought-for-memory; and/or (c) rebuild or regenerate what was apparently forgotten by capitalizing on the conceptual/ideational scaffolding supplied by the understanding. Understanding also provides organizational coherence, thereby chunking and integrating the information into fewer knowledge-representation/retrieval structures, and decreasing the memory burden.

Qualitative explanations, whether direct or analogical, can enhance retention by functioning as an instructional strategy to help a learner establish a meaningful relationship between new information and what is already known and understood. Evidence supporting this approach is provided by Mayer (1975), who used linear and structural explanations on a computer to enhance the learning and retention of software programming. There is considerable evidence to suggest that qualitative explanations help a learner develop models for making new knowledge fit more meaningfully into his or her existing knowledge structures by relating the new knowledge or skill to what has been previously learned and understood. The resulting effect is that the learning and retention of principle- and rule-based complex tasks are enhanced (Gentner, 1980, 1982; Kieras, 1981; Sturges et al., 1981; Tourangeau and Sternberg, 1982).

In summary, qualitative explanations that promote understanding of a to-be-learned task are effective for enhancing retention, presumably because they enable a trainee to reach a higher level of original learning. (The role of such explanations is discussed further in the next chapter as an important part of training techniques that use the expert as a model to guide the trainee.)

Involving the Learner

An important generalization that emerges from several domains of basic and applied research is that long-term retention is enhanced when a learner is an active participant rather than a passive observer during the training process. The inefficiency of humans as passive receivers of information reflects a fundamental property of human memory: a person does not behave like a simple recording apparatus. The storage of new information is a matter of actively interpreting that new information is terms of what is already known, and the reliable retrieval of information from memory requires practicing the retrieval process. Several kinds of research support this generalization.

Cooperative Learning, Peer Teaching, and Related Techniques In the field of education there is abundant evidence that participation by students accomplishes more learning than presentations by instructors. Cooperative learning procedures—in which small groups of students work together on a common problem or project—have been shown to enhance later performance (see, e.g., Johnson et al., 1981; Slavin, 1983) as do peer teaching, proctoring, and coaching programs, in which students participate in the teaching process on a one-on-one or group level (see, e.g., Goldschmid and Goldschmid, 1976; Kulik et al., 1980). In fact, the performance of both the givers and receivers of such student instruction appears to profit from the interaction, and the attitudes of students about the instructional process become more positive as well. In sum, cooperative learning techniques, peer-teaching techniques, and all other such programs in which students take an active role in their own learning lead to improved performance (see, e.g., Rothkopf, 1981).

Practice on Procedural Tasks Overall, it is probably an understatement to say that most training programs involve too much in the way of talking, presenting, and demonstrating on the part of a trainer and too little in the way of answering, producing, and practicing on the part of the trainee. Especially in the case of procedural tasks, listening and watching are ineffective compared with doing—although, of course, doing requires some initial level of learning. For example, watching someone demonstrate how to use an oxygen mask or inflatable lifevest or how to give cardiopulmonary resuscitation is not good preparation for executing those tasks when they are needed. Procedural skills must be practiced and exercised (see, e.g., Schneider, 1985). With increasing complexity of a task, a large amount of practice may be necessary to meet criterion levels of skill, and, initially, components of the task may need to be practiced separately.

The Effects of Generation The "generation effect" (Slamecka and Graf, 1978) refers to the fact that verbal information generated by subjects (learners) in response to cues presented by an experimenter is better remembered at a later time than is information presented for study. Generation effects have been demonstrated with many types of verbal materials and with a variety of initial cuing procedures. It has also been shown (e.g., Wittrock and Carter, 1975) that subjects who generate their own organization of verbal materials (such as a hierarchical grouping of related words) remember those materials better at a later time than do subjects who are simply given such an organization. In general, there is much to be said for the "Socratic method" of instruction, in which the instructor's goal is to get the learner to produce answers and solutions.

Tests as Learning Events Related to the effects of generation is the finding that the act of retrieving information presented earlier facilitates later retrieval of that information (see, e.g., Bjork, 1975; Landauer and Bjork, 1978; Rea and Modigliani, 1985). That is, an act of recall is itself a potent learning event—more potent, in general, than is an opportunity to study the information in question. From the standpoint of long-term retention, tests appear to play the important role of reducing the forgetting that would otherwise take place. Landauer and Ainslie (1975), for example, found that in a college-style technical course, performance on a repeated final examination 1 year later was greatly facilitated by an intervening test with that exam at 6 months. The group with the 6-month test showed virtually no memory loss across the year (from the end-of-course final exam to the repeated final exam 1 year later), and they performed at a much higher level on the 1-year test than did the group with no intervening test.

Another virtue of tests is that they appear to trigger subsequent study opportunities (e.g., Izawa, 1970). That is, more learning appears to take place on the basis of information presented after a test of a learner's memory for that information than takes place without such a test. One interpretation of such results is that tests provide feedback to the learner—clarifying to some extent what has been learned and what remains to be learned—which puts the learner in a better position to take advantage of subsequent information.

The advantage of tests embedded in the training process may grow as the posttraining retention interval gets longer. In contrasting the effects of prior study and test trials, Hogan and Kintsch (1971) found that study trials were superior to test trials in terms of performance at the end of the experimental session, but that test trials were superior to study trials on a test of recall 48 hours later. Once again, then, a condition that may appear to produce better performance during training (in this instance, study trials) may not be optimal in terms of long-term retention.

Refresher Training

Since the focus of this chapter is on original training and what can be done therein to enhance long-term retention and transfer, a thorough discussion of an important related matter—the kinds of posttraining interventions that are useful in maintaining performance at a high level over time—is beyond its scope. In this section, however, we do want to make clear that in many cases posttraining refresher programs are necessary, whatever the program of initial training.

As a retention interval (i.e., nonuse period) increases, the absolute amount of forgetting increases at a negatively accelerated rate (Annett, 1979; Gardlin and Sitterley, 1972; Hagman and Rose, 1983; Hurlock and Montague, 1982; Naylor and Briggs, 1961; Prophet, 1976; Schendel et al., 1978). Refresher learning, practice, or rehearsal is typically needed during such nonuse periods to maintain a given level of knowledge or skill. An important practical consideration, however, is whether a relearning program is even feasible. If it is not, as in emergency situations in which the originally learned knowledge or skill must be at a high enough level to ensure errorless performance in the first real-world execution, then a more intensive refresher program is recommended.

Being able to predict how much forgetting is likely to occur over any given posttraining retention interval is important for determining the refresher training, practice, or rehearsal conditions needed for maintaining performance at an acceptable level. Two approaches to predicting such forgetting have been proposed. An algorithm used by Rigg and Gray (1981; also, Rigg, 1983), which is based on mathematical learning theory, uses performance data from a trainee's initial learning trial to predict the likely rate of forgetting. The approach seems promising, but the empirical tests of the algorithm are too sketchy at present for evaluating it. The algorithm developed by Rose et al. (1984, 1985) for procedural tasks is based on the internal organization or cohesiveness of a given task. This technique, referred to as user's decision aid, estimates how often refresher training should be given to a unit of soldiers to maintain proficiency at a particular level. Although this aid is a promising algorithmic technique for dealing with procedural-task forgetting during nonuse periods, it is obviously more desirable to protect against such forgetting in the first place, if possible, by implementing conditions in training that lead to a high level of original learning and enhanced retention.

The spacing of refresher training, practice, or rehearsal has been found to be important for maintaining any given level of achieved knowledge or skill over nonuse periods. Bahrick (1979) recommends that the spacing be at intervals approximately equal to the expected nonuse interval separating successive occasions when that knowledge or skill needs to be exercised. A related idea is that novices in a given field should get rehearsals of a given skill at about the same intervals that professionals in the field tend to need to exercise that skill.

An important practical point on refresher training is that the practice or rehearsal needs of retrainees appear to be different from those of new trainees. Relatively efficient, cost-effective techniques can be used to maintain a given level of original learning or mastery during nonuse periods for retrainees. One such technique is to substitute covert (imaginary) practice and symbolic rehearsal for the conditions that were used in original training (Annett, 1979; Naylor and Briggs, 1961). Research by Landauer and Bjork (1978) and Bjork (1988) suggests that it may even be optimal to expand the successive intervals between practice sessions. Another technique is to reduce the fidelity of simulation during refresher training, practice, or rehearsal because it does not have to be as high as in original training (Naylor and Briggs, 1961). A third technique is to employ selected part-task training or conceptual simulation instead of the complete original training conditions (Hutchins et al., 1985; Stevens and Steinberg, 1981; Young, 1983). Lastly, brief or partial cuing conditions can also be used as an effective technique: for example, procedural skills can often be rapidly remembered by reminder information, written or oral.

TRANSFER OF TRAINING

As just noted, training performance (i.e., level of original learning) may or may not be an effective predictor of posttraining performance when the training and posttraining contexts are the same or quite similar. When the training and posttraining contexts differ, however, many of the most effective procedures for facilitating the kind of learning that supports transfer apparently impair performance during training. In this section we first discuss the role of level of original learning and perceived similarity between tasks as general factors in the transfer of training; we then discuss some specific procedures during training that enhance transfer to different posttraining contexts.

General Factors in Transfer

Level of Original Learning

The level of original learning is not only a major determinant of retention, it is also a major determinant of transfer. Positive transfer increases with the level of original learning as long as structurally similar responses are required in the training and transfer tasks. The greater the similarity between the tasks, in terms of both stimulus and response requirements, the greater the positive transfer between them (e.g., Ellis, 1965; Osgood, 1949; Schmidt and Young, 1987).

One might expect, for example, some positive transfer between a tennis serve and an overhand volleyball serve because of the similarities in stimuli (a tossed ball) and response requirements (an overhand throwing motion). However, when a new response is paired with a previously learned stimulus, negative transfer may initially occur (Siipola, 1941). An example of negative transfer occurred in the evolution of the butterfly stroke in swimming (described by Fischman et al., 1982). Before the 1960s, the butterfly stroke was swum using a breaststroke kick. The introduction of the dolphin kick produced some negative transfer among butterfly swimmers, probably because of the pairing of this new kick with the traditional armstroke. When there is little or no association between the stimulus-response requirements of two tasks, no transfer is expected: one would not, for example, expect significant transfer effects between the movement patterns of golf and bowling.

It seems important to know the specific relationships between level of original learning, task similarity, and positive transfer, but we found no recent studies that examined the transfer of cognitive or motor tasks as a function of the amount of learning. Several studies of complex problem solving, however, suggest that performance improves with practice of the rules defining the task (e.g., Anzai and Simon, 1979; Kotovsky et al., 1985).

When negative transfer is expected from a training task to a posttraining task—that is, when structurally dissimilar responses are required in the training and transfer tasks—the effects of level of original learning are more complicated. Research on animal learning and on human verbal learning (Mandler, 1968) found that as the level of original learning increases, transfer becomes increasingly negative but that transfer becomes positive at high levels of original learning. Mandler proposed that negative transfer due to response competition increases monotonically to an asymptote as training is extended and the level of original learning increases. Extended training, however, also produces a type of generalized learning that is consistent with the transfer task as well as the training task. Such generalized learning has positive influences that eventually become stronger than the negative influence of response competition. Such an interpretation is consistent with the "learning to learn" idea—that is, learning general problem-solving strategies that are suitable for both training and transfer tasks.

Perceived Similarity Between Tasks

It has been known for some time that, in general, the basis for transfer from a training task to a transfer task are the common components shared by both tasks (Thorndike, 1903). The greater the number of components common to the training and transfer tasks, the greater their similarity, which should lead to greater positive transfer. Gick and Holyoak (1987) propose that *any* salient similarity between training and posttraining tasks will influence a trainee's perceived similarity between the tasks, which in turn will trigger retrieval of the trainee's mental representation of the training situation during the transfer task. The greater the perceived similarity between the training and posttraining tasks, the more likely it is that a trainee will attempt to transfer what was learned during training to the posttraining task. If transfer is attempted, then the direction of transfer—whether it is positive or negative—will be determined by the similarity between the training and posttraining tasks in terms of features that are causally relevant to the goals of the tasks or to the responses required in the posttraining task.

The features of a given task may be either structural or surface characteristics

(Holyoak, 1985). Structural characteristics are causally related to goal attainment; surface characteristics have no such relationship. In the absence of structural similarity, perceived similarity could be based on surface features that may produce negative transfer. It is important to understand that perceived similarity is not simply a function of the objective properties of the two situations. Perceived similarity is also a function of other factors, such as the knowledge or expertise of the individual (Chi et al., 1981) and the context of the two situations (Tversky and Gati, 1978).

It has also been hypothesized that similarity of goals and processing between training and transfer tasks may enhance perceived similarity between those tasks, which will then prompt reminding and transfer even in the absence of explicit instructions or hints to apply the pertinent relevant knowledge (Gick and Holyoak, 1987). In an interesting two-phase experiment, Weisberg et al. (1978) found that the absence of shared goals does not aid memory. In the first phase, subjects were asked to learn a list of paired associates, one of which was the pair box-candle. In the second phase, the subjects were asked to solve Duncker's box-candle problem, which is to figure out a way to attach a candle to a wall using only the materials provided (which include a candle and a box of tacks). The solution involves emptying the box of tacks, using one tack to mount the empty box on the wall, and using melted wax to affix the candle to the top of the box. Weisberg et al. (1978) found that the prior paired-associate learning task did not aid the subsequent problem-solving task. Thus, it is the goal of the task itself, not the particular parts of the problem, that prompts past experience. What is important is the similarity of processing between training and transfer tasks. The chance of obtaining positive transfer is more likely when performers process the training and transfer tasks in a similar way so that compatible responses in both tasks are produced (Bransford and Franks, 1976; Bransford et al., 1979; Lung and Dominowski, 1985; Morris et al., 1977).

In summary, whether or not transfer occurs from training to posttraining is a function of perceived similarity between the two contexts. Perceived similarity of the two tasks is a function of any salient shared component and of a number of other factors, such as expertise and context. To a great extent, an individual's expertise on the subject determines whether the similarities observed are surface features or structural features. Whether actual transfer is positive or negative depends on the actual amount of structural similarity. Transfer is positive when the training and transfer responses are highly similar, that is, when they contain many shared structural components and few distinctive components. Thus, the *amount* of transfer obtained between situations is a function of the perceived similarity; the *direction* of transfer is a function of the objective structural similarity. The greater the perceived similarity of the situations, the greater the amount of transfer. No transfer takes place when two situations are perceived as unrelated, regardless of the degree of response similarity. If a learner does not perceive the similarity between training and posttraining contexts, the level of performance achieved in training clearly will not predict posttraining performance.

Enhancing Transfer

Some of the most promising methods of training for transfer to altered contexts create difficulties for a learner during training. Some of the most promising of those methods involve creating certain types of interference, introducing variability, and reducing the frequency of external feedback.

Providing Contextual Interference During Training

Research on contextual interference shows that learning that requires more cognitive processing is related to better retention and transfer. Contextual interference involves changes in the training context, including changes in the task, practice conditions, and the processing used by trainees. Contending with such changes demands cognitive processing that can, in turn, enhance the level of original learning. These changes have been referred to as "contextual variety," which Battig (1979) believes is closely related conceptually to "transfer appropriate processing" (Bransford et al., 1979). Battig also considers contextual variety to be a way of overcoming the boundaries in memory performance imposed by encoding specificity—in which there is a need to reinstate the original encoding context during a test—to improve performance (Tulving and Thomson, 1973).

One example of contextual interference is provided in a study conducted by Shea and Morgan (1979). They studied the learning of three similar procedural motor tasks that required adult subjects to knock down a series of barriers (with their hands) in a designated order as fast as possible without making any errors. Each of the tasks consisted of a separate pattern of barrier contacts for each trial of training. The three tasks were practiced in two ways, "blocked" and "random": blocked practice involved performing 18 trials of one task before performing 18 trials of each of the other two tasks; random practice involved a random ordering of the three tasks over the 54 total trials. Following training, subjects transferred to either blocked or random conditions with a 10-minute and a 10-day interval. The major finding was that random practice produced poorer performance than blocked practice in training, but it produced superior performance in the posttraining context. This finding has been supported by other studies of adult learners (see Magill and Hall, 1990, for a review).

Another example, using a laboratory task, involved learning certain finger movements; the interference was having or not having the trainees also learn to articulate nonsense terms (e.g., XENF) to match the finger movements. The result was more proficient transfer performance on a different version of the finger task for those who learned the nonsense terms (Battig, 1956, 1966). This finding could be interpreted as showing that the contextual interference between word pronunciations and finger movements generated enhanced transfer. In other words, intratask interference in training produces greater intertask transfer.

Battig (1972, 1979) and others (for reviews, see, e.g., Fendrich et al., 1988; Magill and Hall, 1990; Shea and Zimny, 1983) have confirmed this effect and demonstrated that having to overcome high contextual interference during training produces poor performance in training but enhanced retention and transfer, for both cognitive and motor tasks. Whether or not contextual interference occurs is largely contingent on the degree to which the posttraining task is the same as the one used in original learning. This result follows directly from the "encoding specificity principle" (Tulving and Thomson, 1973) and from the findings of classical intertask transfer research—that an appreciable change from the original learning or encoding situations usually results in decrements in retention or transfer (Battig, 1979).

Battig (1979) explains this effect by proposing that multiple and variable processing strategies have to be used to overcome high contextual interference in order to encode the knowledge or skill being learned. Theoretically, training or practicing under a condition of high contextual interference produces more elaborate and distinctive processing, which enhances retention. Presumably, elaborations during processing produce memory structures for the knowledge or skills learned that are richer and more discriminable and thus lead to easier retrieval. The extent to which positive transfer is enhanced is believed to be a function of the degree to which the contextual interference induces processing strategies that are appropriate for learning other tasks (Morris et al., 1977; Bransford et al., 1979). Since contextual variety should lead to more elaborate and distinctive encoding, it is likely to offer stronger resistance to the typically negative effects that are found when posttraining tasks are changed. In other words, encoding specificity is more likely to be overcome if the original encodings occurred under high contextual variety. Conversely, similar task contexts should induce processing consisting mainly of the development of discriminative and organizational change suited to the specific task demands. Thus, item similarity in original learning of a task should produce better retention or transfer when the task is the same in posttraining as it was in original training. In summary, incorporating contextual variety in training introduces functional interference that makes learning less context dependent and involves trainees in processing activity that in turn produces enhanced retrieval from memory and the ability to adapt their performance to different contexts.

Several studies involving retention and comprehension of verbal materials bear an interesting relationship to the idea of contextual interference. Mannes and Kintsch (1987) had subjects study a brief technical article (on industrial uses of microbes) after studying either an outline that was consistent with the organizational structure of the article or one that was inconsistent (but contained the same information). On tests of verbatim knowledge (e.g., verbatim recall of statements from the article or true-false judgments of whether a given statement did or did not appear in the article), the consistent outline produced better performance, but on tests that required drawing inferences or proposing (or ranking) possible solutions to potential problems, the inconsistent outline produced better performance. The consistent outline apparently resulted in a simpler, more coherent representation in memory, which served to guide verbatim recall and recognition, but did not support processes of generalization and inference. The inconsistent outline, in forcing subjects to resolve organizational discrepancies, apparently resulted in a more embellished, abstracted representation in memory, a representation that contained fewer literal details from the article but that permitted more of the kind of extrapolation that underlies generalization and inference.

Smith et al. (1978) demonstrate that simply varying the environmental context across study sessions on a list of words improves later recall in a novel setting. Smith and Rothkopf (1984) show that such environmental variation can enhance retention of instructional content as well: they manipulated whether four successive 2-hour lectures in a miniature statistics course were given in the same location or in four different locations. They found that recall of key concepts a week after the course was better for students in the varied-context condition. They also found—consistent with the advantages of distributed practice and for both the same-context or varied-context conditions—that presenting the four lectures on four successive days resulted in better recall than presenting all the lectures on one day.

It is of interest to note that training under high contextual interference enhances task retention and transfer in a fashion analogous to supplementary practice: that is, it increases the level of original learning or overlearning. This retention and transfer outcome suggests that training under high contextual interference may be conceptualized as being functionally equivalent to training with additional practice (Farr, 1987). It also suggests that the level of overlearning is being indirectly manipulated by varying the level of contextual interference during training, and so could be viewed as similar to directly manipulating the level of overlearning by varying the amount of supplementary practice.

It can also be argued that even direct manipulations enhance retention and transfer to some extent as a result of the quality of processing during original learning and overlearning. Typically, the enhancement of retention that results from giving supplementary practice to increase the level of overlearning is attributed to the strengthening of connections due directly to the additional trials. However, Mandler (1968) argues against this interpretation because he believes that a retrievable trace is primarily the product of organization—that mere acts of rehearsal or repetition associated with additional practice do not by themselves produce a retrievable trace. Rehearsal or repetition simply allows a learner to establish initial categories and place items into them or to reorganize the categories. Of course, rehearsal or repetition may also provide a learner with the opportunity for more elaborate, deeper processing of information, as Battig (1979) originally proposed. Mandler's position is clearly too extreme in that certain types of stimulus-driven learning take place as a product of mere exposure and repetition, without intentionality on the part of the learner (see Roediger, 1990, for a review of such phenomena); in terms of subsequent purposeful recall of information, however, it seems safe to say that the quality of processing is clearly more important than the duration of processing. The large body of work carried out within the "levels of processing" framework (Craik and Lockhart, 1972) supports this generalization, and that work also supports the generalization that it is the nature of initial processing, not the subject's intent to learn or remember, by itself, that determines later recall performance.

Increasing Variability and Variety in Training

In this section we examine some transfer studies in which performers learned by examples without being given a rule for or a definition of a category. In two of these studies (Gick and Holyoak, 1983; Homa and Cultice, 1984) positive transfer of knowledge was found to increase with the number of examples provided during training for a category-learning task. Theoretically, increases in the number of examples in training should increase the chance of learning the most appropriate rules for transfer on the basis of the features that are structurally associated with category membership. However, optimal positive transfer depends on the representativeness or variability of the examples provided with respect to the category (Anderson et al., 1979). High variability of examples in training generally facilitates transfer; low variability increases the chance that trainees will undertake the transfer task without rules for classifying examples that they have never encountered before (Fried and Holyoak, 1984). Although high variability of examples in training tends to enhance transfer if learning is successful, it may impair learning at the outset, especially if the examples are few in number (Peterson et al., 1973).

When a large variety of examples are given during training, the order in which they are presented may affect learning (Nitsch, 1977). Moreover, the optimal order of examples may depend on performers' approach to the learning task. Using a classification task, Elio and Anderson (1984) found that transfer was better when examples with low-variability were presented first, but only when learners were instructed not to look for a deterministic rule for category membership (i.e., an implicit learning strategy). Examples with high variability examples were best presented first if learners were instructed to look for a rule defining category membership. In general, it is important to realize that instructions given to learners can influence the processing strategies they use during training, which in turn can affect both learning and transfer (Brooks, 1978; McAndrews and Moskovitch, 1985; Medin and Smith, 1981; Sweller et al., 1982).

The findings discussed thus far are based on transfer studies in which subjects learned by examples without being provided with a rule for or definition of a category. When an abstract rule or definition of a category for the task is given to subjects, the inclusion of examples with such abstract training facilitates transfer (Cheng et al., 1986; Gick and Holyoak, 1983; Nitsch, 1977). There is some evidence, however, that in certain situations only a minimal number of examples need to be provided (Fong et al., 1986). It appears that including an abstract rule with examples in training is particularly useful when the training and transfer items are superficially dissimilar (Gick and Holyoak, 1983) or when it is difficult to determine the rule from the examples alone (Cheng et al., 1986; Fong et al., 1986).

In the motor domain, much of the recent research on variability of practice has been conducted within Schmidt's (1975, 1982) schema theory of discrete motor skill learning (discussed above). This theory proposes that practice produces abstract rules that control classes of movement responses, with each class being represented by a generalized motor program. For example, kicking motions such as those involved in kicking a soccer ball are assumed to be generated by a generalized motor program. Kicking a soccer ball a certain distance at a given speed is produced by specifying parameters (e.g., overall duration or force of the response) for the generalized program. Parameters are selected on the basis of schemata or rules, developed from past experience with the program: these schemata specify the association between the environmental outcomes of the kicking movements and the values of the parameters selected. When an individual wants to kick a certain distance, his or her schema specifies the parameter for the generalized motor program for kicking, and the program is executed with this parameter value. Thus, the movements needed to generate each individual kick do not have to be stored or represented, and the soccer player can produce novel kicking movements that have not been used before. An important prediction of this theory is that increased variability of practice pertinent to the generalized program would impair performance in training, but would yield enhanced performance in posttraining on a transfer test to a novel task within the same response class. Many studies have investigated this prediction, and although the evidence is not entirely consistent, there is a reasonable amount of support for it (e.g., Catalano and Kleiner, 1984; Margolis and Christina, 1981; for a review, see Shapiro and Schmidt, 1982).

Recently, however, Schmidt and Young (1987) have interpreted the variability of practice effects found in previous studies in terms of blocked or random practice, that is, in terms of contextual interference. They suggest that the variability-of-practice effects may be nothing more than random-practice effects. Certain results, however, obtained by Wulf and Schmidt (1988) with adults and by Pigott and Shapiro (1984) and Wrisberg and Mead (1983) with children, seem difficult to interpret in terms of contextual interference effects.

In summary, training manipulations that involve number, variability, and order of examples for category-learning tasks and variability of practice for motor-learning tasks may impair training performance, but they appear to increase learning so that posttraining performance is enhanced. This is yet another line of evidence that indicates that the level of performance achieved in training is not a good predictor of learning and posttraining performance in transfer.

Reducing Feedback

Feedback is an integral part of most training programs. It can be defined as information resulting from an action, and it can be intrinsic or augmented. Intrinsic feedback is the information people receive as a natural consequence of their actions and takes two forms: proprioceptive feedback is the sensing of muscle, joint, or tendon activity, as in the feeling from executing a golf swing; visual feedback comes from the outcome of an activity, as in the flight path of a golf ball. Augmented feedback, which is the focus of this section, is information performers would not ordinarily receive as a result of their actions. It is provided by a source external to the performers, such as an instructor, mirror, or videotape system. Augmented feedback can be verbal, as when an instructor explains how to correct an error, or it can be nonverbal, as when an instructor demonstrates how to make a correction or shows a videotape replay of an individual's performance. Typically, augmented feedback comparing the performance outcome with some goal outcome is referred to as knowledge of results; augmented feedback that characterizes the movement pattern itself is referred to as knowledge of performance. (For other descriptions of different types of feedback, see e.g., Schmidt, 1988:423-426).

It has been known for some time that augmented feedback has a substantial effect on learning and performance during training (e.g., Thorndike, 1927). The focus here is on the influences that feedback has on retention and transfer during training. Much of the recent relevant research has been carried out in the motor domain (for reviews, see Newell, in press; Salmoni et al., 1984; Schmidt, in press). That research has challenged a commonly accepted generalization about augmented feedback that emerged from an abundance of prior research (see Bilodeau, 1966, 1969). That generalization is that any increase in feedback in training—in its immediacy, accuracy, or frequency—will improve learning and performance. Over the years, that generalization served as a basis for incorporating augmented feedback in the design of training programs and simulators. However, evidence from some recent studies in the motor domain (e.g., Schmidt et al., 1989; Winstein, 1988; Winstein and Schmidt, 1990) and several earlier studies in the verbal domain (e.g., Landauer and Bjork, 1978; Krumboltz and Weisman, 1962) raise questions about the validity of the generalization.

Generally, these studies manipulated augmented feedback in training by giving it less frequently, such as on every fifth trial instead of every trial; or by giving it on every early trial but gradually eliminating it on later trials; or by giving it in summary form over a set of trials. Essentially, these studies reveal that training with augmented feedback that is given less frequently or in summary form produces poorer performance in training than feedback administered after every trial, but it produces better posttraining performance in retention and transfer. These results can be interpreted as indicating that frequent augmented feedback during training functions primarily to guide behavior toward the criterion (i.e., training) performance but that it also may create a dependency in which the feedback is relied on to guide behavior, and the learning needed to produce proficient posttraining performance in retention or transfer either does not occur or occurs only at a weak level. This dependency on frequent augmented feedback could be the result of the trainees' not using information-processing strategies that would ordinarily be used to learn the task in training if feedback were available less often. Without having adequately learned the task, the trainees are at a disadvantage for retention or transfer when there is no (or less) feedback.

In summary, augmented feedback has traditionally been structured to bring about rapid acquisition of a task so that some criterion level of performance in training is achieved as quickly as possible. But the evidence now suggests that some of the commonly accepted ways in which augmented feedback has been manipulated to facilitate training performance are less than optimal for enhancing learning and posttraining performance. The evidence in this section (see also Schmidt and Bjork, 1992) indicates that the kind of feedback manipulations that enhance learning and posttraining performance actually decreases the rate at which performance improves during training.

The common denominator of the training procedures reviewed above is that they teach processes that can be called on by a posttraining task at a later time, particularly if the posttraining task and setting differ from the training task and setting. That is, such procedures induce “transfer-appropriate processing” (Morris et al., 1977). In responding to the “difficulties” introduced by contextual interference, variability in the conditions of practice, reduced feedback, and so forth, the learner is taught to carry out processes that result in a more elaborated mental representation of the task—a representation that can, to some extent, be used in a different context. The learner is better prepared, so to speak, not only to perceive the similarities between the training task and the different versions of that task in posttraining contexts, but also better equipped to perform by having achieved the more generalized declarative and procedural knowledge demanded by that category of task.

CONCLUSIONS AND IMPLICATIONS FOR TRAINING

Measuring Learning and Performance The effectiveness of a training program should be measured not by the speed of acquisition of a task during training or by the level of performance reached at the end of training, but, rather, by a learner's performance in the posttraining tasks and real-world settings that are the target of training.

Two important dimensions of posttraining performance are the ability to resist forgetting and interference over periods of disuse of a given skill and the ability to generalize training to contexts and tasks that differ in their surface characteristics from the training contexts or tasks. Depending on the relative priorities given to those two dimensions of posttraining performance, the optimal package of training components will differ somewhat. One general principle, however, is that tests of a learner's progress during training should, as much as possible, measure performance as it will be measured on the posttraining task(s) in the posttraining setting(s).

Retention Given posttraining tasks and conditions that are identical or similar to the training tasks and conditions, posttraining performance is enhanced as the level of original learning is increased. That level can be increased by putting greater demands on the learner—making the criterion of mastery more difficult, for example, or requiring supplementary (postmastery) practice after the criterion has been reached. Introducing variations in the conditions and sequencing of practice, the immediate consequence of which is to degrade performance, may be a particularly promising way to increase the level of original learning, and, hence, posttraining retention.

Skills that demand little attention or effort to perform are regarded as automatic; the more automatic a given skill, the higher the likelihood that the skill can be retained over nonuse periods without refresher training. Certain types of procedural tasks, however, tend to be easily forgotten, especially when their components have a low degree of internal organization or cohesiveness. The rate of forgetting of procedural tasks is a function of the number of steps needed to perform the task, and the steps most likely to be forgotten are those not cued by the equipment, environment, or preceding steps.

Several instructional strategies to enhance the retention and transfer of procedural tasks can be derived from the research on learning: relating the knowledge to be learned to the relevant knowledge learners already have in memory; teaching techniques (e.g., mnemonics) that learners can use to provide their own elaborations; having the training regimen require repeated use of the knowledge to be learned; and providing for and encouraging the use and elaboration of acquired knowledge and skill during nonuse periods. In general, a learner should be an active participant, not a passive observer, during the training process.

However well designed the initial training, refresher training may still be needed during posttraining periods of disuse in order to maintain a given level of knowledge and skill; refresher training can become less frequent over time. The training needs of retrainees are different from those of new trainees; relatively efficient, cost-effective techniques can be used to maintain a given level of original learning in retrainees.

Transfer of Training In general, the similarity of goals and cognitive processing between training and transfer tasks is a critical factor in enhancing transfer. A learner, therefore, should be challenged by means of manipulation of practice variables, such as feedback, contextual interference, and number and variability of examples. These manipulations, which may impair training performance, not only help the learner to process the learning task more deeply, but also suggest appropriate processes for transfer, particularly to related but distinct posttraining tasks.

REFERENCES

Anderson, J.R., P.J. Kline, and C.M. Beasley 1979 A general learning theory and its application to schema abstraction. In G.H. Bower, ed., *The Psychology of Learning and Motivation,* Vol. 13. New York: Academic Press.

Annett, J. 1979 Memory for skill. In M.M. Gruneberg and P.E. Morris, eds., *Applied Problems in Memory.* London: Academic Press.

Anzai, Y., and H.A. Simon 1979 The theory of learning by doing. *Psychological Review* 86:124-140.

Baddeley, A.H., and D.J.A. Longman 1978 The influence of length and frequency on training sessions on the rate of learning to type. *Ergonomics* 21:627-635.

Bahrick, H.P. 1979 Maintenance of knowledge: questions about memory we forgot to ask. *Journal of Experimental Psychology, General* 108:296-308.

1987 Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 13:344-349.

Battig, W.F. 1956 Transfer from verbal pretraining to motor performance as a function of motor task complexity. *Journal of Experimental Psychology* 51:371-378.

1966 Facilitation and interference. In E.A. Bilodeau, ed., *Acquisition of Skill.* New York: Academic Press.

1972 Intratask interferences as a source of facilitation in transfer and retention. In R.F. Thompson and J.F. Voss, eds., *Topics in Learning and Performance.* New York: Academic Press.

1979 The flexibility of human memory. In L.S. Cermark and F.I.M. Craik, eds., *Levels of Processing in Human Memory.* Hillsdale, N.J.: Erlbaum.

Bell, H.M. 1950 Retention of pursuit rotor task after one year. *Journal of Experimental Psychology* 40:648-649.

Bilodeau, I.M. 1966 Information feedback. In E.A. Bilodeau, ed., *Acquisition of Skill.* New York: Academic Press.

1969 Information feedback. In E.A. Bilodeau, ed., *Principles of Skill Acquisition.* New York: Academic Press.

Bjork, R.A. 1975 Retrieval as a memory modifier: an interpretation of negative recency and related

phenomena. In R.L. Solso, ed., *Information Processing and Cognition.* New York: Wiley.
1979 Information-processing analysis of college teaching. *Educational Psychologist* 14: 15-23.
1988 Retrieval practice and the maintenance of knowledge. Pp. 397-401 in M.M. Gruneberg, P.E. Morris, and R.N. Sykes, eds., *Practical Aspects of Memory II.* London: Wiley.
Bransford, J.D., and J.J. Franks 1976 Toward a framework for understanding learning. In G.H. Bower, ed., *The Psychology of Learning and Motivation,* Vol. 10. New York: Academic Press.
Bransford, J.D., J.J. Franks, C.D. Morris, and B.S. Stein 1979 Some general constraints on learning and memory research. In L.S. Cermack and F.I.M. Craik, eds., *Levels of Processing in Human Memory.* Hillsdale, N.J.: Erlbaum.
Brooks, L. 1978 Nonanalytic concept formation and memory for instances. In E. Rosch and B. Lloyd, eds., *Cognition and Categorization.* Hillsdale, N.J.: Erlbaum.
Catalano, J.R., and B.M. Kleiner 1984 Distant transfer and practice variability. *Perceptual and Motor Skills* 58:851-856.
Cheng, P.W., K.J. Holyoak, R.E. Nisbett, and L.M. Oliver 1986 Pragmatic versus syntactic approaches to training deductive reasoning. *Cognitive Psychology* 18:293-328.
Chi, M., P. Feltovich, and R. Glaser 1981 Categorization and representation of physics problems by experts and novices. *Cognitive Science* 5:121-152.
Craik, F.I.M., and R.S. Lockhart 1972 Levels of processing: a framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11:671-684.
Crowder, R.G. 1976 *Principles of Learning and Memory.* Hillsdale, N.J.: Erlbaum.
Dempster, F.N. 1990 The spacing effect: a case study in the failure to apply the results of psychological research. *American Psychologist* 43:627-634.
Ebbinghaus, H. 1913 *Memory* (H.A. Ruger and C.E. Bussenius, trans.). New York: Teachers College. (Original work published 1885; paperback ed., New York: Dover, 1964).
Elio, R., and J. Anderson 1984 The effects of information order and learning mode on schema abstraction. *Memory and Cognition* 7:397-417.
Ellis, H.C. 1965 *The Transfer of Learning.* New York: Macmillan.
Farr, M.J. 1987 *The Long-Term Retention of Knowledge and Skills: A Cognitive and Instructional Perspective.* New York: Springer-Verlag.
Fendrich, D.W., A.F. Healy, L. Meiskey, R.J. Crutcher, W. Little, and L.E. Bourne, Jr. 1988 Skill Maintenance: Literature Review and Theoretical Analysis. Technical report AFHRL-TP-87-73. Air Force Human Resources Laboratory, Brooks Air Force Base, Texas.
Fischman, M.G., R.W. Christina, and M.J. Vercruyssen 1982 Retention and transfer of motor skills: a review for the practitioner. *Quest* 33:181-194.
Fong, G.T., D.H. Krantz, and R.E. Nisbett 1986 The effects of statistical training on thinking about everyday problems. *Cognitive Psychology* 18:253-292.

Fried, L.S., and K.J. Holyoak 1984 Induction of category distributions: a framework for classification learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10:234-257.

Gardlin, G.R., and T.E. Sitterly 1972 Degradation of Learned Skills: A Review and Annotated Bibliography. Boeing Company, Seattle, Washington.

Gentner, D. 1980 The Structure of Analogical Models in Science. Technical Report 4451. Bolt Beranek and Newman, Inc., Cambridge, Massachusetts.

1982 Are scientific analogies metaphors? In D.S. Miall, ed., *Metaphor: Problems and Perspectives.* Brighton, England: Harvester Press, Ltd.

Gentner, D., and A.L. Stevens 1983 *Mental Models.* Hillsdale, N.J.: Erlbaum.

Gick, M.L., and K.J. Holyoak 1983 Schema induction and analogical transfer. *Cognitive Psychology* 15:1-38.

1987 The cognitive basis of knowledge transfer. In S.M. Cormier and J.D. Hagman, eds., *Transfer of Learning: Contemporary Research and Applications.* San Diego, Calif.: Academic Press.

Glenberg, A.M. 1979 Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory and Cognition* 7:95-112.

Glenberg, A.M., and T.S. Lehmann 1980 Spacing repetitions over 1 week. *Memory and Cognition* 8:528-538.

Goldschmid, B., and M.L. Goldschmid 1976 Peer teaching in higher education: a review. *Higher Education* 5:9-33.

Hagman, J.D., and A.M. Rose 1983 Retention of military skills: a review. *Human Factors* 25:199-213.

Hogan, R.M., and W. Kintsch 1971 Differential effects of study and test trials on long-term recognition and recall. *Journal of Verbal Learning and Verbal Behavior* 10:562-567.

Holyoak, K.J. 1985 The pragmatic of analogical transfer. In G.H. Bower, ed., *The Psychology of Learning and Motivation,* Vol. 19. New York: Academic Press.

Homa, D., and J. Cultice 1984 Role of feedback, category size, and stimulus distortion on the acquisition and utilization of ill-defined categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10:83-94.

Horton, D.L., and C.B. Mills 1984 Human learning and memory. *Annual Review of Psychology* 35:361-394.

Hurlock, R.E., and W.E. Montague 1982 Skill Retention and Its Implications for Navy Tasks: An Analytic Review. NPRDC SR 82-21. Navy Personnel Research and Development Center, San Diego, Calif.

Hutchins, E.L., J.D. Hollan, and D.A. Norman 1985 Direct Manipulation Interfaces. ICS Report 8503. University of California, La Jolla.

Izawa, C. 1970 Optimal potentiating effects and forgetting-prevention effects of tests in paired-associate learning. *Journal of Experimental Psychology* 83:340-344.

Johnson, D.W., G. Maruyama, R. Johnson, and D. Nelson 1981 Effects of cooperative, competitive, and individualistic goal structures on achievement: a meta analysis. *Psychological Bulletin* 89:47-62.

Jones, M.B. 1985 Nonimposed Overpractice and Skill Retention. Technical report no. 86-55. Army Research Institute for the Behavioral and Social Sciences, Alexandria, Va.

Jonides, J., M. Naveh-Benjamin, and J. Palmer 1985 Assessing automaticity. *Acta Psychological* 60:157-171.

Kieras, D.E. 1981 Knowledge Representation in Cognitive Psychology. Technical report no. 7. Personnel and Training Research Programs, Office of Naval Research, Arlington, Va.

Kieras, D.E., and S. Boviar 1984 The role of mental model in learning to operate a device. *Cognitive Science* 8:255-273.

Konoske, P.J., and J.A. Ellis 1985 Cognitive Factors in Learning and Retention of Procedural Tasks. Paper presented at the April meeting of the American Education Research Association, Chicago, Illinois.

Kotovsky, K., J.R. Hayes, and H.A. Simon 1985 Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology* 17:248-294.

Krumboltz, J.D., and R.G. Weisman 1962 The effect of intermittent confirmation in programmed instruction. *Journal of Educational Psychology* 53:250-253.

Kulik, J.A., C.C. Kulik, and P.A. Cohen 1980 Effectiveness of computer-based college teaching: a meta-analysis of findings. *Review of Educational Research* 50:525-544.

Landauer, T.K., and K.I. Ainslie 1975 Exams and use as preservatives of course-acquired knowledge. *The Journal of Educational Research* 69(3):99-105.

Landauer, T.K., and R.A. Bjork 1978 Optimum rehearsal patterns and name learning. In M.M. Gruneberg, P.E. Morris, and R.N. Sykes, eds., *Practical Aspects of Memory.* London: Academic Press.

Lee, T.D., and E.D. Genovese 1988 Distribution of practice in motor skill acquisition: learning and performance effects reconsidered. *Research Quarterly for Exercise and Sport* 59:277-287.

Loftus, G.R. 1985 Evaluating forgetting curves. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11:397-406.

Lung, C.T., and R. Dominowski 1985 Effects of strategy instructions and practice on nine-dot problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 11:804-811.

Magill, R.A., and K.G. Hall 1990 A review of the contextual interference effect in motor skill acquisition. *Human Movement Science* 9:241-289.

Mandler, G. 1968 Association and organization: facts, fancies, and theories. In T.R. Dixon and D.L. Horton, eds., *Verbal Behavior and General Behavior Theory.* Englewood Cliffs, N.J.: Prentice-Hall.

Mannes, S.M., and W. Kintsch 1987 Knowledge organization and text organization. *Cognition and Instruction* 4:91-115.

Margolis, J., and R.W. Christina 1981 A test of Schmidt's schema theory of discrete motor skill learning. *Research Quarterly for Exercise and Sport* 52:474-483.

Mayer, R.E. 1975 Different problem-solving competencies established in learning computer programming with and without meaningful models. *Journal of Educational Psychology* 67:725-734.

McAndrews, M.P., and M. Moskovitch 1985 Rule-based and exemplar-based classification in artificial grammar learning. *Memory and Cognition* 13:469-475.

McGeoch, J.A., and A.L. Irion 1952 *The Psychology of Human Learning,* 2nd ed. New York: Longman, Green, and Company.

Medin, D.L., and E.E. Smith 1981 Strategies and classification learning. *Journal of Experimental Psychology: Human Learning and Memory* 7:241-253.

Melnick, M.J. 1971 Effects of overlearning on the retention of a gross motor skill. *Research Quarterly* 42:60-69.

Melton, A.W. 1970 The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior* 9:596-606.

Montague, W.E. 1988 What Works: Summary of Research Findings with Implications for Navy Instruction and Learning. Technical report NAVEDTRA 115-1. Office of the Chief of Naval Education and Training, Pensacola, Florida.

Morris, C.D., J.D. Bransford, and J.J. Franks 1977 Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Verbal Behavior* 16:519-533.

Morris, C.D., B.S. Stein, and J.D. Bransford 1979 Prerequisites for the utilization of knowledge in recall of prose passages. *Journal of Experimental Psychology: Human Learning and Memory* 5:253-261.

Naylor, J.C., and G.E. Briggs 1961 Long-Term Retention of Learned Skill: A Review of the Literature. ASD technical report 61-390. Laboratory of Aviation Psychology, Ohio State University.

Newell, K.M. In press Augmented information and the acquisition of skill. In R. Daugs and K. Blischke, eds., *Motor Learning and Training.* Schorndorf, West Germany: Hoffman.

Nitsch, K.E. 1977 Structuring Decontextualized Forms of Knowledge. Unpublished doctoral dissertation. Vanderbilt University.

Osgood, C.E. 1949 The similarity paradox in human learning: a resolution. *Psychological Review* 56:132-143.

Owens, J., G.H. Bower, J.B. Black 1979 The "soap opera" effect in story recall. *Memory and Cognition* 7:185-191.

Peterson, M.J., R.B. Meagher, Jr., H. Chait, and S. Gillie 1973 The abstraction and generalization of dot patterns. *Cognitive Psychology* 4:378-398.

Pigott, R.E., and D.C. Shapiro 1984 Motor schema: the structure of the variability session. *Research Quarterly for Exercise and Sport* 55:41-45.

Prophet, W.W. 1976 Long-Term Retention of Flying Skills: A Review of the Literature. HumRRO final report 76-35. Human Resources Research Organization, Alexandria, Va.

Rea, C.P., and V. Modigliani 1985 The effect of expanded versus massed practice on the retention of multiplication facts and spelling lists. *Human Learning* 4:11-18.

Rigg, K.E. 1983 Optimization of Skill Retention in the U.S. Army Through Initial Training and Analysis and Design. McFann-Gray and Associates, Monterey, Calif.

Rigg, E.E., and B.B. Gray 1981 Estimating Skill Training and Retention Functions Through Instructional Model Analysis. McFann-Gray and Associates, Monterey, Calif.

Roediger, H.L., III 1990 Implicit memory: retention without remembering. *American Psychologist* 45:1043-1056.

Rose, A.M., M.Y. Czarnolewski, F.E. Gragg, S.H. Austin, P. Ford, J. Doyle, and J.D. Hagman, Jr. 1984 Acquisition and Retention of Soldering Skills. Technical report no. 671. Army Research Institute for the Behavioral and Social Sciences, Alexandria, Va.

Rose, A.M., P.H. Radtke, H.H. Shettel, and J.D. Hagman 1985 User's Manual for Predicting Military Task Retention. Report no. AIR FR37800. American Institutes for Research, Washington, D.C.

Rothkopf, E.Z. 1981 A macroscopic model of instruction and purposeful learning. *Instructional Science* 10:105-122.

Rothkopf, E.Z., and E.V. Coke 1966 Variations in phrasing and repetition interval and the recall of sentence materials. *Journal of Verbal Learning and Verbal Behavior* 5:86-91.

Salmoni, A.W., R.A. Schmidt, and C.B. Walter 1984 Knowledge of results and motor learning: a review and critical reappraisal. *Psychological Bulletin* 95:355-386.

Schendel, J., J. Shields, and M. Katz 1978 Retention of Motor Skills: Review. Technical Paper 313. U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, Va.

Schendel, J.D., and J.D. Hagman 1982 On sustaining procedural skills over a prolonged retention interval. *Journal of Applied Psychology* 67:605-610.

Schmidt, R.A. 1975 A schema theory of discrete motor skill learning. *Psychological Review* 82:225-260.

1982 The schema concept. Pp. 219-235 in J.A.S. Kelso, ed., *Human Motor Behavior: An Introduction.* Hillsdale, N.J.: Erlbaum.

1988 *Motor Control and Learning: A Behavioral Emphasis,* 2nd ed. Champaign, Ill.: Human Kinetics.

In press Frequent augmented feedback can degrade learning: evidence and interpretations. In G.E. Stelmach and J. Requin, eds., *Tutorials in Motor Neuroscience.* Dordrecht, The Netherlands: Kluwer.

Schmidt, R.A., and R.A. Bjork 1992 New conceptualizations of practice: common principles in three research paradigms suggest important new concepts for practice. *Psychological Science* (June).

Schmidt, R.A., and D.E. Young 1987 Transfer of movement control in motor skill learning. In S.M. Cormier and J.D. Hagman, eds., *Transfer of Learning: Contemporary Research and Applications.* San Diego, Calif.: Academic Press.

Schmidt, R.A., D.E. Young, S. Swinnen, and D.C. Shapiro 1989 Summary knowledge of results for skill acquisition: support for the guidance by pothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15:352-359.

Schneider, W. 1985 Training high-performance skills: fallacies and guidelines. *Human Factors* 27:285-300.

Schneider, W., and R.M. Shiffrin 1977 Controlled and automatic information processing: detection, search, and attention. *Psychological Review* 84:1-66.

Schneider, W., S.T. Dumais, and R.M. Shiffrin 1984 Automatic and control processing and attention. In R. Parasuraman and D.R. Davies, eds., *Varieties of Attention.* Orlando, Fla.: Academic Press.

Shapiro, D.C., and R.A. Schmidt 1982 The schema theory: recent evidence and developmental implications. In J.A.S. Kelso and J.E. Clark, eds.,*The Development of Movement Control and Coordination.* New York: Wiley.

Shea, J.B., and R.L. Morgan 1979 Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory* 5:179-187.

Shea, J.B., and S.T. Zimny 1983 Context effects in memory and learning movement information,. In R.A. Magill, ed., *Memory and Control of Action.* Amsterdam, Holland: North-Holland.

Shields, J.L., S.L. Goldberg, and J.D. Dressel 1979 Retention of Basic Soldering Skills. Research report no. 1225. U.S. Army Research Institute for the Behavioral and Social Sciences, Alexandria, Va.

Shiffrin, R.M., and W. Schneider 1977 Controlled and automatic human information processing: perceptual learning, automatic attending, and a general theory. *Psychological Review* 84:127-190.

Siipola, E.M. 1941 The relation of transfer to similarity in habit-structure. *Journal of Experimental Psychology* 28:233-261.

Slamecka, M.J., and P. Graf 1978 The generation effect: delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning and Memory* 4:592-604.

Slamecka, M.J., and B. McElree 1983 Normal forgetting of verbal lists as a function of their degree of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 9:384-397.

Slavin, R.E. 1983 *Cooperative Learning.* New York: Longman, Inc.

Smith, E., and L. Goodman 1984 Understanding instructions: the role of explanatory material. *Cognition and Instruction* 1:359-396.

Smith, S.M., and E.Z. Rothkopf 1984 Contextual enrichment and distribution of practice in the classroom. *Cognition and Instruction* 1(3):341-358.

Smith, S.M., A.X. Glenberg, and R.A. Bjork 1978 Environmental context and human memory. *Memory and Cognition* 6:342-353.

Stevens, A., and C. Steinberg 1981 Project Steamer: Taxonomy for Generating Explanations of How to Operate Complex Physical Devices. NPRDC Technical note 81-21. Navy Personnel Research and Development Center, San Diego, Calif.

Sturges, P., J. Ellis, and W. Wulfeck 1981 Effects of Performance-Oriented Text Upon Long-Term Retention of Factual Mate rial. NPRDC Technical report no. 81-22). Navy Personnel Research and Development Center, San Diego, Calif.

Sweller, J., R. Mawer, and W. Howe 1982 Consequences of history-cued and means-ends strategies in problem solving. *American Journal of Psychology* 95:455-483.

Thorndike, E.L. 1927 The law of effect. *American Journal of Psychology* 39:212-222.

1903 *Educational Psychology.* New York: Lemcke and Buechner.

Tourangeau, R., and R. Sternberg 1982 Understanding and appreciating metaphors. *Cognition* 11:203-244.

Tulving, E., and D.M. Thomson 1973 Encoding specificity and retrieval processes in episodic memory. *Psychological Review* 80:352-373.

Tulving, T. 1985 How many memory systems are there? *American Psychologist* 40:385-398.

Tversky, A., and I. Gati 1978 Studies in similarity. In E. Rosch and B.B. Lloyd, eds., *Cognition and Categorization.* Hillsdale, N.J.: Erlbaum.

Walberg, H.J. 1990 Productive teaching and instruction: assessing the knowledge base. *Phi Delta Kappa* 71:470-478.

Weisberg, R., M. Di Camillo, and D. Phillips 1978 Transferring old associations to new situations: a non-automatic process. *Journal of Verbal Learning and Verbal Behavior* 17:219-228.

Wertheim, A.H. 1985 Some Remarks on the Retention of Learned Skills. Paper presented at NATO Conference on Transfer of Training, September-October, Brussels, Belgium.

Wetzel, S.K., P.J. Konoske, and W.E. Montague 1983 Estimating Skill Degradation for Aviation and Antisubmarine Warfare Operators (AWs): Loss of Skill and Knowledge Following Training. Technical report NPRDC SR-83-31. Navy Personnel Research and Development Center, San Diego, Calif.

Winograd, T.W. 1975 Frame representations and the declarative-procedural controversy. In D.G. Bobrow and A.M. Collins, eds., *Representation and Understanding: Studies in Cognitive Science.* New York: Academic Press.

Winstein, C.J. 1988 Relative Frequency of Information Feedback in Motor Performance and Learning. Unpublished doctoral dissertation, University of California, Los Angeles.

Winstein, C.J., and R.A. Schmidt 1990 Reduced frequency of knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16:677-691.

Wittrock, M.C., and J.F. Carter 1975 Generative processing of hierarchically organized words. *American Journal of Psychology* 88:489-501.

Wrisberg, C.A., and B.J. Mead 1983 Developing coincident timing skill in children: a comparison in training methods. *Research Quarterly for Exercise and Sport* 54:67-74.

Wulf, G., and R.A. Schmidt 1988 Variability in practice: facilitation in retention and transfer through schema formation or context effects? *Journal of Motor Behavior* 20:133-149.

Young, R.M. 1983 Surrogates and mappings: two kinds of conceptual models for interactive devices. In D. Gentner and A. Stevens, eds., *Mental Models.* Hillsdale, N.J.: Erlbaum.