

# CHAPTER 35

## Selection, Training, and Development of Personnel

**ROBERT W. SWEZEY**

InterScience America, Inc.

**RICHARD B. PEARLSTEIN**

American Red Cross

<b>1. INTRODUCTION</b>	<b>920</b>	3.6. Acquisition, Retention, and Transfer of Learned Information	929
<b>2. SELECTION</b>	<b>921</b>	3.6.1. Acquisition	929
2.1. Predictors	921	3.6.2. Retention	930
2.1.1. Aptitude and Ability Tests	921	3.6.3. Transfer	931
2.1.2. Simulations	922	3.7. Training Teams and Groups	933
2.1.3. Interviews and Biodata	922	3.8. Training Evaluation	934
2.1.4. Work Samples	923	<b>4. DEVELOPMENT</b>	<b>937</b>
2.1.5. References	923	4.1. Individual Performance Enhancement	937
2.1.6. Others	923	4.2. Organizational Performance Enhancement	938
2.2. Validity and Fairness	923	4.2.1. Management and Leadership Development	938
2.3. Performance Measurement and Criterion Assessment	924	4.2.2. Organizational Development	939
<b>3. TRAINING</b>	<b>924</b>	<b>5. THE FUTURE OF SELECTION, TRAINING, AND DEVELOPMENT</b>	<b>939</b>
3.1. Differences between Training and Education	924	5.1. Selection	939
3.2. Training, Learning, Performance, and Outcomes	924	5.2. Training	940
3.3. Analysis of Organizations, Jobs, Tasks, and People	925	5.3. Development	940
3.4. Training Design and Development	926	<b>REFERENCES</b>	<b>941</b>
3.5. Training Techniques, Strategies, and Technologies	928		
3.5.1. Simulation as a Training Technology	929		

### 1. INTRODUCTION

Selection, training, and development are the primary tools used to promote human performance in industrial engineering settings. In this chapter, we will review important research issues in these areas. We have elected to emphasize training over the equally important topics of selection and development, primarily because training can often have substantial impact in existing application areas. In many industrial engineering situations, staff members are already in place, and even if personnel selection is an option, it is performed by other agencies, such as the human resources

department. Second, “development” activities often consist of training—either the construction of new training programs or the use of existing ones, or both. Consequently, in our view, training (plus judicious use of other, ancillary, performance-improvement technologies) constitutes a serious potential impact area for industrial engineering situations.

Further, a major concern in the domain of selection research is the development of technologies designed to address the validity and reliability of instruments used to select personnel. (“Validity” deals with the question of determining the extent to which an assessment technique actually measures what it purports to measure, such as intelligence, clerical skill, and mathematical aptitude. “Reliability” addresses the extent to which such an instrument measures whatever it purports to measure consistently, that is, without major score differences upon repeated applications of the instrument to the same people.) Such topics as these, while very important, are beyond the scope of this chapter.

For the same reason, we have chosen not to deal with other widely researched topics in this domain, such as recruitment of personnel, employee retention, personnel turnover, metaanalysis (statistical technologies for collapsing data from multiple research programs in order to reach generalizable conclusions), and validity generalization (techniques designed to apply known measurement validities obtained in one situation to another). Readers interested in these topics may consult the many current textbooks available in these areas or an overall review article (Guion 1991). An authoritative reference for those involved in the methods used in selection and classification of personnel is the *Standards for Educational and Psychological Testing* (AERA et al. 1999).

## 2. SELECTION

In selection research, *predictor* refers to any technique or technology used to predict the success of individuals in the application context. In general, it is the goal of selection technologies to predict successfully how well people will perform if hired or selected for admission to a specific situation, such as a college. In the following pages, we briefly review several of the most popular predictor categories.

### 2.1. Predictors

#### 2.1.1. *Aptitude and Ability Tests*

Very generally, *aptitude* refers to a capacity or suitability for *learning* in a particular domain or topic area (such as mechanical aptitude), whereas *ability* refers to a natural or inborn talent to *perform* in a given area (such as athletic ability). Development and application of tests of aptitude and ability have long been a major focus of activity in the area of selection research. In general, ability and aptitude tests consist of several subcategories. Cognitive ability tests address issues that involve remembering information, producing ideas, thinking, and comparing or evaluating data. Sensory motor tests include physical, perceptual, motor, and kinesiologic measures. Personality, temperament, and motivational tests address those issues as predictors of job success. Additionally, so-called lie-detector or polygraph tests are increasingly used in some situations as a selection and/or screening device, despite the existence of a substantial amount of data that question their validity (see Guion 1991; Sackett and Decker 1979).

One series of projects that have recently addressed aptitude testing involves the development and validation of the (U.S.) Armed Services Vocational Aptitude Battery (ASVAB) (Welsh et al. 1990). This effort involved years of research and dozens of studies that applied and researched the contributions of aptitude-testing techniques to a huge variety of military job settings. To transfer these military technologies to civilian settings, researchers compared ASVAB counterparts to the civilian General Aptitude Test Battery (GATB). These comparisons were reviewed by Hunter (1986) and others, who have demonstrated that many ASVAB components are highly similar to their civilian counterparts, thus arguing that many of the military’s scales and testing procedures are directly relevant for nonmilitary applications.

A second major research program in this area is known as Project A. That effort (see Campbell 1990 for a review) involved a number of complex research projects designed to validate predictors of hands-on job-performance measures and other criteria for U.S. Army jobs. Project A, conducted at a cost of approximately \$25 million, is certainly one of the largest selection-research projects in history. It found that “core job performance” was predicted by general cognitive ability (assessed by the ASVAB), while ratings of effort and leadership, personal discipline, and physical fitness were best predicted by personality measures (McHenry et al. 1990). Other predictors produced only small increments in validity over general cognitive ability in predicting core job performance (Schmidt et al. 1992).

A common issue in selection research is the question of whether “general” ability is more important in determining job success than are various “special” abilities (which presumably are components of the more overall “general” factor). Various authors (Hunter 1986; Jensen 1986; Thorndike 1986; and many others) have argued in favor of the general factor, whereas others (cf. Prediger 1989)

have challenged these conclusions, suggesting that specific aptitudes are more important in predicting job performance.

### 2.1.2. *Simulations*

Simulations are increasingly used as predictors in selection situations. Simulations are representations of job situations. Although presumably task relevant, they are nonetheless abstractions of real-world job tasks. Many simulations used in selection occur in so-called assessment centers, in which multiple techniques are used to judge candidates. In such situations, people are often appraised in groups so that both interpersonal variables and individual factors can be directly considered. Simulation exercises are often used to elicit behaviors considered relevant to particular jobs (or tasks that comprise the jobs) for which people are being selected. This includes so-called in-basket simulations for managerial positions, where hypothetical memos and corporate agenda matters are provided to job candidates in order to assess their decision-making (and other) skills.

Recently, computer-based simulations—ranging from low-fidelity “games” to highly sophisticated complex environments coupled with detailed performance-measurement systems—have been developed for use as predictors in selection and assessment situations. One example of this approach is team performance assessment technology (TPAT) (e.g., Swezey et al. 1999). This and similar methods present simulated task environments that assess how individuals and/or teams develop plans and strategies and adapt to changes in fluctuating task demands. This technology exposes respondents to complex environments that generate task-relevant challenges. At known times, participants have an opportunity to engage in strategic planning; at other times, emergencies may require decisive action. These technologies are designed to allow participants to function across a broad range of situations that include unique demands, incomplete information, and rapid change. They often employ a performance-measurement technology termed “quasi-experimental” (Streufert and Swezey 1985). Here, fixed events of importance require that each participant deal with precise issues under identical conditions. Other events, however, are directly influenced by actions of participants. In order to enable reliable measurement, fixed (preprogrammed) features are inserted to allow for comparisons among participants and for performance assessment against predetermined criteria. Other methodologies of this sort include the tactical naval decision making system (TANDEM), a low-fidelity simulation of a command, control, and communication environment (Dwyer et al. 1992; Weaver et al. 1993), and the team interactive decision exercise for teams incorporating distributed expertise (TIDE<sup>2</sup>) (Hollenbeck et al. 1991).

### 2.1.3. *Interviews and Biodata*

Most reviews concerning the reliability and validity of interviews as a selection device “have ended with the depressing but persistent conclusion that they have neither” (Guion 1991, p. 347). Then why are they used at all? Guion provides four basic reasons for using interviews as selection devices in employment situations:

1. They serve a public relations role. Even if rejected, a potential employee who was interviewed in a competent, professional manner may leave with (and convey to others) the impression that he was treated fairly.
2. They can be useful for gathering ancillary information about a candidate. Although collecting some forms of ancillary information (such as personal, family, and social relationship data) is illegal, other (perfectly legitimate) information on such topics as work history and education that may have otherwise been unavailable can be collected in a well-structured interview protocol.
3. They can be used to measure applicant characteristics that may otherwise not be adequately measured, such as friendliness, ability to make a good first impression, and conversational etc.
4. Interviewers are themselves decision makers. In some circumstances, interviewers may be used to arrive at overall judgments about candidate suitability that can be used in the decision process.

Many alternative ways of conducting interviews have been studied. For instance, Martin and Nagao (1989) compared paper-and-pencil, face-to-face, and computerized versions of interviews. They found that socially desirable responses often occurred when impersonal modes of interviewing were used, but that applicants for high-status jobs often resented impersonal interviews. Tullar (1989) found that successful applicants are given longer interviews and dominate the conversation more than less-successful applicants.

As to biodata, information about an individual’s past, Hammer and Kleiman (1988) found that only 6.8% of 248 firms surveyed stated that they had ever used biodata in employment decisions, and only 0.4% indicated that they currently used biodata. Yet there is evidence that biodata has

substantial validity as a selection device (Schmidt et al. 1992). The most frequent reasons for not using biodata were (a) lack of knowledge, (b) lack of methodological expertise, and (c) lack of personnel, funding, and time.

An important assumption in use of biodata as a selection tool is that past behaviors—often behaviors far in the past (e.g., high school achievements)—are good predictors of future behaviors. However Kleiman and Faley (1990) found that biodata items inquiring about *present* behaviors were just as valid as those focusing on past behaviors in predicting intention to reenlist in the Air National Guard; and Russell et al. (1990) found that retrospective life-history essays could serve as the source of valid biodata. These and other studies indicate that biodata may be a significantly underutilized source of information in personnel selection situations.

#### **2.1.4. Work Samples**

Unlike simulations, which are not the actual activities performed in the workplace but an abstraction or representation of aspects of those activities, work sample testing involves the actual activities performed on the job. These are measured or rated for use as a predictor. The logic applied to work sample testing is somewhat different from other selection technologies in that it employs a “criterion-based” perspective: the performance being tested *is* usually a subset of the actual task employed and is derived directly from the content domain of that task. This stands in contrast to other types of selection methodologies where a measure of some “construct” (such as personality, aptitude, ability, or intelligence) is measured and related statistically (usually correlated) with actual job performance (see Swezey 1981 for a detailed discussion of this issue).

In work sample testing, a portion of an actual clerical task that occurs everyday may be used to measure clerical performance (e.g., “type this letter to Mr. Elliott”). Typically, such tests actually sample the work domain in some representative fashion; thus, typing a letter would be only one aspect of a clerical work sample test that may also include filing, telephone answering, and calculating tasks.

Work sample testing is less abstract than other predictive techniques in that actual performance measurement indices (e.g., counts, errors) or ratings (including checklists) may be used as indices of performance. Where possible, work sample testing (or its close cousin, simulation) is the preferred method for use in selection research and practice because one does not get mired down trying to decide how to apply abstract mental constructs to the prediction of job performance.

#### **2.1.5. References**

The use of references (statements from knowledgeable parties about the qualifications of a candidate for a particular position) as a predictor of job performance or success is the most underresearched—yet widely used—selection technique. The problem with using references as a selection technique is that, in most cases, the persons supplying the references have unspecified criteria. That is, unless the requesting parties provide explicit standards for use in providing the reference, referees may address whatever qualities, issues, or topics they wish in providing their comments. This wide-ranging, subjective, unstructured technique is therefore of highly questionable validity as a predictor of job performance. Nevertheless, the technique is widely used in employment and other application-oriented settings (e.g., college applications) because we typically value the unstructured opinions of knowledgeable people concerning assets and/or liabilities of an applicant or candidate. The key word here is “knowledgeable.” Unfortunately, we often have no way of determining a referee’s level of knowledge about a candidate’s merits.

#### **2.1.6. Others**

Among the many other predictors used as selection devices are such areas as graphology, genetic testing, investigative background checks, and preemployment drug testing. Of these, preemployment drug testing has undoubtedly received the most attention in the literature. In one noteworthy effort, Fraser and Kroeck (1989) suggested that the value of drug screening in selection situations may be very low. In partial reaction to this hypothesis, a number of investigations were reviewed by Schmidt et al. (1992). Those authors cite the work of Fraser and Kroeck and of Murphy et al. (1990), who, using college student samples, found that individual drug use and self-reported frequency of drug use were consistently correlated with disapproval of employee drug testing. However, no substantial associations between acceptability of drug testing and employment experience or qualifications were established.

### **2.2. Validity and Fairness**

As mentioned above, test validity is a complex and ambiguous topic. Basically, one can break this domain into two primary categories: (1) psychometric validities, including criterion-based validities (which address—either in the future [predictive] or simultaneously [concurrent]—the relationship(s) among predictors and criteria), and construct validity (which relates predictors to psychological con-

structs, such as abilities, achievement, etc.); and (2) content (or job/task-related) validities. See Guion (1991) for a discussion of these topics.

Various U.S. court rulings have mandated that job-related validities be considered in establishing test fairness for most selection and classification situations. These issues are addressed in the latest *Standards for Educational and Psychological Testing* (1999) document.

### 2.3. Performance Measurement and Criterion Assessment

An oft-quoted principle is that the best predictor of future performance is past performance. However, current performance can be a good predictor as well. In selecting personnel to perform specific tasks, one may look at measurements of their performance on similar tasks in the past (e.g., performance appraisals, certification results, interviews, samples of past work). As noted earlier, one may also measure their current performance on the tasks for which one is selecting them, either by assessing them on a simulation of those tasks or, when feasible, by using work samples. A combination of both types of measures may be most effective. Measures of past performance show not only what personnel can do, but what they have done. In other words, past performance reflects motivation as well as skills. A disadvantage is that, because past performance occurred under different circumstances than those under which the candidates will perform in the future, ability to generalize from past measures is necessarily limited. Because measures of current performance show what candidates can do under the actual (or simulated) conditions in which they will be performing, using both types of measures provides a more thorough basis on which to predict actual behaviors.

Whether based on past performance or current performance, predictors must not only be accurate, they must also be relevant to the actual tasks to be performed. In other words, the criteria against which candidates' performance is assessed must match those of the tasks to be performed (Swezey 1981). One way to ensure that predictors match tasks is to use criterion-referenced assessment of predictor variables for employee selection. To do this, one must ensure that the predictors match the tasks for which one is selecting candidates as closely as possible on three relevant dimensions: the conditions under which the tasks will be performed, the actions required to perform the task, and the standards against which successful task performance will be measured. While it is better to have multiple predictors for each important job task (at least one measure of past performance and one of current performance), it may not always be possible to obtain multiple predictors. We do not recommend using a single predictor for multiple tasks—the predictor is likely to be neither reliable nor valid.

## 3. TRAINING

### 3.1. Differences between Training and Education

“Training” (the systematic, structured development of specific skills required to perform job tasks) differs from “education” (development of broad-based informational background and general skills). Most people are more familiar with education than training because (a) their experiences from elementary school through college usually fall within an educational model, and (b) much of what is called training in business and industry is actually more akin to education than training.

Education is important, even essential, for building broad skill areas, but training is the approach of preference for preparing people to perform specific tasks or jobs. Table 1 shows differences between education and training and between educators and trainers. Not all differences apply to any specific instance of training or education, but as a whole they illustrate the distinction between the two.

### 3.2. Training, Learning, Performance, and Outcomes

The *purpose* of training is to facilitate learning skills and knowledges required to perform specific tasks. The *outcome* of training is acceptable performance on those tasks. People often learn to perform tasks without training, but if training is effective, it enables people to learn faster and better. For example, a person might be able to learn to apply the principle of cantilevering to bridge construction through reading and trial and error, but the process would be lengthier and the outcomes would probably be less precise than if the person received systematic, well-designed training.

Because training is a subsystem of a larger organizational system, it should not be treated outside the context of the larger system. In organizations, training is usually requested in two types of circumstances: either current performance is inadequate or new performance is required. In other words, organizations train either to solve current problems or to take advantage of new opportunities. A second factor affecting training is whether it is intended for new or incumbent employees. Basically, training for new opportunities is essentially the same for new and incumbent employees. However, new employees may require prerequisite or remedial training prior to receiving training on new tasks to help take advantage of opportunities, whereas incumbents presumably would have already completed the prerequisite training (or would have otherwise mastered the necessary skills).

**TABLE 1 Education and Educators Compared to Training and Trainers**

Education	Training
Increases knowledge	Increases job performance
Broadens understanding	Focuses on job skills
Conceptualizes experience	Simulates experience
Asks "How much do you know?"	Asks "What can you do with what you know?"
Provides base of knowledge	Uses knowledge to build skills
Often requires broad generalization of knowledge	Requires knowledge to be applied in specific situations
Educator	Trainer
Dispenses information	Facilitates learning experience
Is primary learning resource	Allows participants to serve as learning resources
Stresses inquiry	Stresses practical application
Groups students according to ability (norm-referenced)	Tries to help all participants perform to specified level (criterion-referenced)
Focuses on information	Focuses on performance
Evaluates via classroom testing	Evaluates via on-the-job performance assessment

Just as learning can take place without training, organizational outcomes can be achieved without training. In fact, when acceptable performance on specific tasks is a desired organizational outcome, training is frequently *not* the appropriate way to achieve that outcome (see Gilbert 1978; Harless 1975; Mager and Pipe 1970). The type of analysis that determines factors contributing to performance of an organization's personnel subsystem has been called by a variety of names, including performance analysis (Mager and Pipe 1970), and front-end analysis (Harless 1975). Mager and Pipe (1970), noting that training is appropriate only when performance is unacceptable because performers lack necessary skills and knowledge, suggested this classic question for making the determination: "Could they do it if their lives depended on it?" If the answer is "no," then training is warranted. If the answer is "yes," the next question is, "What is preventing them from performing as required?" Mager and Pipe also suggested various ways of categorizing obstacles to desired performance outcomes that do not consist of skill/knowledge deficiencies (see also Harless 1975; Rummier and Brache 1990). These can be reduced to two broad categories:

1. Lack of environmental support (e.g., well-specified processes, job aids, tools and equipment, management expectations, adequate facilities)
2. Lack of appropriate motivational systems (e.g., clear and immediate feedback, appropriate consequences for action, well-designed reward systems)

For skill/knowledge deficiencies, training interventions are appropriate. Deficiencies involving environmental support and/or appropriate motivational systems should be remedied through other human performance interventions, such as providing clear instructions and developing reward systems. Regardless of the intervention, implementation should always include pilot testing and evaluation components. If desired outcomes are not achieved, the analysis, design, and/or implementation may have been flawed.

### 3.3. Analysis of Organizations, Jobs, Tasks, and People

Organizations consist of people performing tasks to achieve a particular mission. A job is defined as a collection of related tasks that helps accomplish significant elements of an organization's products and/or services. For training and other human performance-improvement interventions to be successful, they must address appropriate organizational levels and/or units. Organizational analysis seeks to determine relationships among an organization's inputs, functions, jobs, and products/services. This type of analysis builds upon front-end and performance analysis.

Rummier and Wilkins (1999) introduced a technique known as performance logic to diagnose an organization's problems and opportunities. The process highlights disconnects among procedures, skills, jobs, and desired product/service outcomes, pointing the way to appropriate training and other

human performance-improvement interventions. Others (Mourier 1998) have also demonstrated successes using this approach.

Job analysis has several purposes. One deals with determining how to apportion an organization's tasks among jobs. This process identifies the tasks that belong with each job, the goal being to reduce overlap while providing clean demarcations between products and processes. From the perspective of an organization's human resources subsystem, job analysis can be used to organize tasks into clusters that take advantage of skill sets common to personnel classification categories (e.g., tool and die press operator, supervisory budget analyst) and consequently reduce the need for training. By distributing tasks appropriately among jobs, an organization may be able to select skilled job candidates, thereby reducing attendant training costs.

A second purpose of job analysis involves job redesign. During the 1970s, researchers addressed various ways to redesign jobs so that the nature of the jobs themselves would promote performance leading to desired organizational outcomes. For example, Lawler et al. (1973) found that several factors, such as variety, autonomy, and task identity, influenced employee satisfaction and improved job performance. Herzberg (1974) identified other factors, such as client relationship, new learning, and unique expertise, having similar effects on satisfaction and performance. Peterson and Duffany (1975) provide a good overview of the job redesign literature.

Needs analysis (sometimes called needs assessment) examines what should be done so that employees can better perform jobs. Needs analysis focuses on outcomes to determine optimal performance for jobs. Rossett (1987) has provided a detailed needs-analysis technique. Instead of needs analysis, many organizations unfortunately conduct a kind of wants analysis, a process that asks employees and/or supervisors to state what is needed to better perform jobs. Because employees and managers frequently cannot distinguish between their wants and their needs, wants analysis typically yields a laundry list of information that is not linked well to performance outcomes.

Finally, task analysis is a technique that determines the inputs, tools, and skills/knowledge necessary for successful task performance. In training situations, task analysis is used to determine the skill/knowledge requirements for personnel to accomplish necessary job outcomes. Skills gaps are defined as the difference between required skills/knowledge and those possessed by the individuals who are (or will be) performing jobs. Training typically focuses on eliminating the skills gaps. Helpful references for conducting task analyses include Carlisle (1986) and Zemke and Kramlinger (1982).

### 3.4. Training Design and Development

A comprehensive reference work with respect to training design and development (Goldstein 1993) refers to what is generally known as a systems approach to training. Among many other things, this approach includes four important components:

1. Feedback is continually used to modify the instructional process, thus training programs are never completely finished products but are always adaptive to information concerning the extent to which programs meet stated objectives.
2. Recognition exists that complicated interactions occur among components of training such as media, individual characteristics of the learner, and instructional strategies.
3. The systems approach to training provides a framework for reference to planning.
4. This view acknowledges that training programs are merely one component of a larger organizational system involving many variables, such as personnel issues, organization issues, and corporate policies.

Thus, in initiating training program development, it is necessary, according to Goldstein, to consider and analyze not only the tasks that comprise the training, but characteristics of the trainees and organizations involved.

This approach to instruction is essentially derived from procedures developed in behavioral psychology. Two generic principles underlie technological development in this area. First, the specific behaviors necessary to perform a task must be precisely described, and second, feedback (reinforcement for action) is utilized to encourage mastery.

This (and previous work) served as a partial basis for the military's instructional systems development (ISD) movement (Branson et al. 1975) and the subsequent refinement of procedures for criterion-referenced measurement technologies (Glaser 1963; Mager 1972; Swezey 1981).

For a discussion of the systems approach to training, we briefly review the Branson et al. (1975) ISD model. This model is possibly the most widely used and comprehensive method of training development. It has been in existence for over 20 years and has revolutionized the design of instruction in many military and civilian contexts. The evolutionary premises behind the model are that performance objectives are developed to address specific behavioral events (identified by task analyses), that criterion tests are developed to address the training objectives, and that training is essentially developed to teach students to pass the tests and thus achieve the requisite criteria.

The ISD method has five basic phases: analysis, design, development, implementation, and control (evaluation).

Phase one—analysis:

- (a) Perform a task analysis. Develop a detailed list of tasks, the conditions under which the tasks are performed, the skills and/or knowledge required for their performance, and the standards that indicate when successful performance has been achieved.
- (b) Select pertinent tasks/functions. Note, by observation of an actual operation, the extent to which each task is actually performed, the frequency with which it is performed, and the percentage of a worker's time devoted to it. Then make an assessment concerning the importance of each task to the overall job.
- (c) Construct job-performance measures from the viewpoint that job-performance requirements are the basis for making decisions about instructional methods.
- (d) Analyze training programs that already exist to determine their usefulness to the program being developed.
- (e) Identify the instructional setting/location for each task selected for instruction.

Phase two—design:

- (a) Describe student entry behavior. Classify, at the task level, categories into which each entry behavior falls (i.e., cognitive, procedural, decision making, problem solving, etc.). Then check whether each activity is within the existing performance repertoire of the trainees (something the potential students already know) or is a specialized behavior (something the students will have to learn).
- (b) Develop statements that translate the performance measures into terminal learning objectives. Consider the behavior to be exhibited, the conditions under which the learning will be demonstrated, and the standards of performance that are considered acceptable.
- (c) Develop criterion-referenced tests (CRTs). CRTs are tests that measure what a person needs to know or do in order to perform a job. Ensure that at least one test item measures each instructional objective. Determine whether the testing should be done by pencil and paper, by practical exercise, or by other methods. Criterion-referenced tests perform three essential functions:
  - They help to construct training by defining detailed performance goals.
  - They aid in diagnosis of whether a student has absorbed the training content because the tests are, in fact, operational definitions of performance objectives.
  - They provide a validation test for the training program.
- (d) Determine the sequence and structure of the training. Develop instructional strategies and the sequencing of instruction for the course.

Phase three—development:

- (a) Specify specific learning events and activities.
- (b) Specify an instructional management plan. Determine how instruction will be delivered and how the student will be guided through it: by group mode, self-paced mode, or combinations of the two. Determine how the curriculum will be presented (e.g., lectures, training aids, simulators, job-performance aids, television, demonstration, and computer based).
- (c) Review and select existing instructional materials for inclusion in the curriculum.
- (d) Develop new instructional materials.
- (e) Validate the instruction by having it reviewed by experts or job incumbents or by trying it out on typical students.

Phase four—implementation:

- (a) Implement the instructional management plan by ensuring that all materials, equipment, and facilities are available and ready for use. Before beginning instruction, schedule students and train instructors.
- (b) Conduct the instruction under the prescribed management plan in the designated setting under actual conditions using final course materials.

Phase five—control (evaluation):

- (a) Monitor course effectiveness, assess student progress, assess the quality of the course, and compare results with the original learning objectives.
- (b) Determine how graduates perform on the job.
- (c) Revise the course as appropriate.



### 3.5. Training Techniques, Strategies, and Technologies

Two factors known to affect quality of training are the techniques used to deliver the instruction and the methods or strategies used to convey the instructional content. One of the main tasks in the development of training, therefore, is the selection and implementation of appropriate instructional methods and techniques.

One function of technology-assessment research is to provide a basis for estimating the utility of particular technology-based approaches for training. Researchers have analyzed virtually every significant technology trend of the last two decades, including effects of programmed instruction (Kulik et al. 1980), computer-based instruction (Kulik and Kulik 1987), and visual based instruction (Cohen et al. 1981). The majority of these studies show no significant differences among technology groups and, when pooled, yield only slight advantages for innovative technologies. Kulik and associates report that the size of these statistically significant gains is in the order of 1.5 percentage points on a final exam. Cohen et al. (1981) conducted an analysis of 65 studies in which student achievement using traditional classroom-based instruction was compared with performance across a variety of video-based instructional media, including films, multimedia, and educational TV. They found that only one in four studies (approximately 26%) reported significant differences favoring visual-based instruction. The overall effect size reported by Cohen et al. was relatively small compared to studies of computer-based instruction or of interactive video.

Computer-based instruction and interactive video both represent areas where some data reporting advantages have emerged; specifically, reductions in the time required to reach threshold performance levels (Fletcher 1990; Kulik and Kulik 1987). Results of an analysis of 28 studies, conducted by Fletcher (1990), suggested that interactive video-based instruction increases achievement by an average of 0.50 standard deviations over conventional instruction (lecture, text, on-the-job training, videotape).

Results of decades of work in the area of media research led Clark (1983, 1994) to conclude that media have no significant influence on learning effectiveness but are mere vehicles for presenting instruction. According to Clark (1983, p. 445), "the best current evidence is that media are mere vehicles that deliver instruction but do not influence student achievement any more than the truck that delivers our groceries causes changes in our nutrition." Similarly, Schramm (1977) commented that "learning seems to be affected more by what is delivered than by the delivery system." Although students may prefer sophisticated and elegant media, learning appears largely unaffected by these features. Basically, all media can deliver either excellent or ineffective instruction. It appears that it is the instructional methods, strategies, and content that facilitate or hinder learning. Thus, many instructional technologies are considered to be equally effective provided that they are capable of dealing with the instructional methods required to achieve intended training objectives.

Meanwhile, teaching methods appear to be important variables influencing the effectiveness of instructional systems. Instructional methods define how the process of instruction occurs: what information is presented, in what level of detail, how the information is organized, how information is used by the learner, and how guidance and feedback are presented. The choice of a particular instructional method often limits the choice of presentation techniques. Technology-selection decisions, therefore, should be guided by: (1) capacity of the technology to accommodate the instructional method, (2) compatibility of the technology with the user environment, and (3) trade-offs that must be made between technology effectiveness and costs.

Conventional instructional strategies such as providing advance organizers (Allen 1973), identifying common errors (Hoban and Van Ormer 1950), and emphasizing critical elements in a demonstration (Travers 1967) work well with many technology forms, including video-based instructional technologies. Embedding practice questions within instructional sequences also appears to be a potent method supported by interactive video and computer-based instruction. Benefits of embedded questions apparently occur whether responses to the questions are overt or covert or whether the material is preceded or followed by the questions (Teather and Marchant 1974). Compared to research on instructional delivery systems, relatively little research has examined the application of instructional methods.

Morris and Rouse (1985) have suggested that strategies used to teach troubleshooting skills should combine both theoretical and procedural knowledge. This was confirmed by Swezey et al. (1991), who addressed these issues using a variety of instructional strategies. Three strategies, termed "procedural," "conceptual," and "integrated," were used. Results indicated that integrated training strategies that combine procedural and theoretical information facilitate retention of concepts and improve transfer to new settings. Swezey and Llaneras (1992) examined the utility of a technology that provides training program developers with guidance for use in selecting *both* instructional strategies and media. The technology distinguishes among 18 instructional strategies, as well as among 14 types of instructional media. Outputs are based upon decisions regarding the types of tasks for which training is to be developed, as well as consideration of various situational characteristics operating in the particular setting. Three general classes of situational characteristics are considered: (1) task

attributes (motion, color, difficulty, etc.), (2) environmental attributes (hazardous, high workload, etc.), and (3) user attributes (level of experience and familiarity with the task). Results indicated that participants who had access to the selection technology made better presentation strategy, media, and study/practice strategy selections than those relying on their own knowledge.

### **3.5.1. *Simulation as a Training Technology***

As technology develops, progress is being made in linking various types of instructional devices. Fletcher (1990) has reviewed much of the literature in this area. In general, he has found that computer-based instructional techniques were often more effective than conventional instruction if they included interactive features such as tutorials. Fletcher also notes that the available studies provide little insight into the relative contributions of various features of interactive technology to learning.

Equipment simulators have been used extensively for training aircraft pilots. Jacobs et al. (1990), and Andrews et al. (1992) have reviewed this research. In general, it appears that simulator training combined with training in actual aircraft is more effective than training in the aircraft by itself. According to Tannenbaum and Yukl (1992), the realism of simulators is being enhanced greatly by continuing developments in video technology, speech simulation, and speech recognition; and advancements in networking technology have opened up new possibilities for large-scale simulator networking. Although simulators have been used extensively for training small teams in the military, networking allows larger groups of trainees to practice their skills in large-scale interactive simulations (Thorpe 1987; Alluisi 1991; Weaver et al. 1995). Often in these circumstances, the opponent is not merely a computer but other teams, and trainees are therefore often highly motivated by the competition. In such situations, as in the actual environment, simulations are designated to operate directly as well as to react to unpredictable developments. A thorough history of the development of one such complex, networked system, known as SIMNET, has been provided by Alluisi (1991). Comprehensive, system-based evaluations of these technologies, however, remain to be performed.

## **3.6. Acquisition, Retention, and Transfer of Learned Information**

### **3.6.1. *Acquisition***

Much literature has indicated that it is necessary to separate the initial acquisition of a performance skill or concept from longer-term learning (Schmidt and Bjork 1992; Cormier 1984; Salmoni et al. 1984). Schmidt and Bjork (1992), for instance, have described a process by which characteristics of a task create an environment in which performance is depressed but learning is apparently enhanced. The implications of this phenomenon, known as contextual interference (Battig 1966; Shea and Morgan 1979), for training are that variations of task conditions, manipulation of task difficulty, or other means of inducing extra effort by trainees are likely to be beneficial for retention and transfer. Recent research by Schneider et al. (1995) investigated effects of contextual interference in rule-based learning, a skill underlying many applied tasks found in the aviation, military, and computer fields. Results mirrored those found in the verbal learning and motor skill-acquisition literatures, indicating that a random practice schedule leads to the best retention performance after initial acquisition but hinders initial learning. Thus, training conditions designed to achieve a particular training objective (long-term retention, transfer, and/or resistance to altered contexts) are not necessarily those that maximize performance during acquisition. This concept has significant implications for predicting trainee performance. Although a typical method employed to predict trainee performance involves monitoring acquisition data during training, using immediate past performance data as predictors to estimate future retention performance, this strategy may not appropriately index performance on the job or outside of the training environment. Further, initial performance of complex skills tends to be unstable and often is a poor indicator of final performance. Correlations between initial and final performance levels for a grammatical reasoning task, for example, reached only 0.31 (Kennedy et al. 1980), a moderate level at best. Using initial performance during learning as a predictor, therefore, may lead to inconsistent and/or erroneous training prescriptions.

Skill acquisition is believed by many to proceed in accordance with a number of stages or phases of improvement. Although the number of stages and their labels differ among researchers, the existence of such stages is supported by a large amount of research and theoretical development during the past 30 years. Traditional research in learning and memory posits a three-stage model for characterizing associative-type learning involving the process of differentiating between various stimuli or classes of stimuli to which responses are required (stimulus discrimination), learning the responses (response learning), and connecting the stimulus with a response (association). Empirical data suggest that this process is most efficient when materials are actively processed in a meaningful manner, rather than merely rehearsed via repetition ( Craik and Lockhart 1972; Cofer et al. 1966).

Anderson (1982) has also proposed a three-stage model of skill acquisition, distinguishing among cognitive, associative, and autonomous stages. The cognitive stage corresponds to early practice in which a learner exerts effort to comprehend the nature of the task and how it should be performed.

In this stage, the learner often works from instructions or an example of how a task is to be performed (i.e., modeling or demonstration). Performance may be characterized by instability and slow growth, or by extremely rapid growth, depending upon task difficulty and degrees of prior experience of the learner. By the end of this stage, learners may have a basic understanding of task requirements, rules, and strategies for successful performance; however, these may not be fully elaborated. During the associative stage, declarative knowledge associated with a domain (e.g., facts, information, background knowledge, and general instruction about a skill acquired during the previous stage) is converted into procedural knowledge, which takes the form of what are called production rules (condition–action pairs). This process is known as knowledge compilation. It has been estimated that hundreds, or even thousands, of such production rules underlie complex skill development (Anderson 1990). Novice and expert performance is believed to be distinguished by the number and quality of production rules. Experts are believed to possess many more elaborate production rules than novices (Larkin 1981).

Similarly, Rumelhart and Norman (1978) recognized three kinds of learning processes: (1) the acquisition of facts in declarative memory (accretion), (2) the initial acquisition of procedures in procedural memory (restructuring), and (3) the modification of existing procedures to enhance reliability and efficiency (tuning). Kyllonen and Alluisi (1987) reviewed these concepts in relation to learning and forgetting facts and skills. Briefly, they suggest that new rules are added to established production systems through the process of accretion, fine-tuned during the process of tuning, and subsequently reorganized into more compact units during the restructuring process.

Rasmussen (1979) has also distinguished among three categories, or modes, of skilled behavior: skill based, rule based, and knowledge based. Skill-based tasks are composed of simple stimulus-response behaviors, which are learned by extensive rehearsal and are highly automated. Rule-based behavior is guided by conscious control and involves the application of appropriate procedures based on unambiguous decision rules. This process involves the ability to recognize specific well-defined situations that call for one rule rather than the other. Knowledge-based skills are used in situations in which familiar cues are absent and clear and definite rules do not always exist. Successful performance involves the discrimination and generalization of rule-based learning. Rasmussen proposes that, in contrast to Anderson's model, performers can move among these modes of performance as dictated by task demands. This general framework is useful in conceptualizing links between task content and the type of training required for proficiency in complex tasks.

### 3.6.2. *Retention*

Instructional designers must consider not only how to achieve more rapid, high-quality training, but also how well the skills taught during training will endure after acquisition. Further, what has been learned must be able to be successfully transferred or applied to a wide variety of tasks and job-specific settings. Swezey and Llaneras (1997) have recently reviewed this area.

Retention of learned material is often characterized as a monotonically decreasing function of the retention interval, falling sharply during the time immediately following acquisition and declining more slowly as additional time passes (Wixted and Ebbesen 1991; Ebbinghaus 1913; McGeoch and Irion 1952). There is general consistency in the loss of material over time. Subjects who have acquired a set of paired items, for example, consistently forget about 20% after a single day and approximately 50% after one week (Underwood 1966). Bahrck (1984), among others, demonstrated that although large parts of acquired knowledge may be lost rapidly, significant portions can also endure for extended intervals, even if not intentionally rehearsed.

Evidence suggests that the rate at which skills and knowledge decay in memory varies as a function of the degree of original learning; decay is slower if material has previously been mastered than if lower-level acquisition criteria were imposed (Loftus 1985). The slope and shape of retention functions also depend upon the specific type of material being tested as well as upon the methods used to assess retention. As meaningfulness of material to the student increases, for example, rate of forgetting appears to slow down. Further, recognition performance may be dramatically better than recall, or vice versa, depending simply upon how subjects are instructed (Tulving and Thomson 1973; Watkins and Tulving 1975). Also, various attributes of the learning environment, such as conditions of reinforcement, characteristics of the training apparatus, and habituation of responses, appear to be forgotten at different rates (Parsons et al. 1973).

Retention, like learning and motivation, cannot be observed directly; it must be inferred from performance following instruction or practice. To date, no uniform measurement system for indexing retention has been adopted. General indices of learning and retention used in the research literature over the past 100 years include a huge variety of methods for measuring recall, relearning, and recognition. Luh (1922), for instance, used five different measures to index retention: recognition, reconstruction, written reproduction, recall by anticipation, and relearning.

The list of variables known to reliably influence rate of forgetting of learned knowledge and skills is relatively short. In a recent review, Farr (1987) surveyed the literature and identified a list of

variables known to influence long-term retention, including degree of original learning, characteristics of the learning task, and the instructional strategies used during initial acquisition. According to Farr, the single largest determinant of magnitude of retention appears to be the degree of original learning. In general, the greater the degree of learning, the slower will be the rate of forgetting (Underwood and Keppel 1963). This relationship is so strong that it has prompted some researchers to argue that any variable that leads to high initial levels of learning will facilitate retention (Prophet 1976; Hurlock and Montague 1982).

The organization and complexity of material to be learned also appear to have a powerful influence on retention. The efficiency of information acquisition, retrieval, and retention appears to depend to a large degree on how well the material has been organized (Cofer et al. 1966). Conceptually organized material appears to show considerably less memory loss and to be more generalizable than material that is not so organized.

Research has identified numerous variables that fall under the banner of strategies for skill and knowledge acquisition and retention, including spaced reviews, massed/distributed practice, part/whole learning, and feedback. An important variable with respect to forgetting is spaced practice. The "spacing effect" (the dependency of retention on the spacing of successive study sessions) suggests that material be studied at widely spaced intervals if retention is required. Similarly, research comparing distributed practice (involving multiple exposures of material over time) vs. massed practice (requiring concentrated exposure in a single session), has occurred in a wide variety of contexts, including the acquisition of skills associated with aircraft carrier landings (Wightman and Sistrunk 1987), word-processing skills (Bouzd and Crawshaw 1987), perceptual skills associated with military aircraft recognition (Jarrard and Wogalter 1992), and learning and retention of second-language vocabulary (Bloom and Shuell 1981). Most research in this area, however, has emphasized acquisition of verbal knowledge and motor skills. In general, research on the issue of distribution of practice has emphasized effects on acquisition and found that distribution of practice is superior to massed practice in most learning situations, long rest periods are superior to short rest periods, and short practice sessions between rests yield better performance than do long practice sessions (Rea and Modigliani 1988). However, no consistent relationships between these variables and long-term retention have emerged.

Techniques that help learners to build mental models that they can use to generate retrieval cues, recognize externally provided cues, and/or generate or reconstruct information have also generally facilitated retention (Kieras and Bovair 1984). Gagné (1978) identified three general strategies for enhancing long-term retention, including reminding learners of currently possessed knowledge that is related to the material to be learned, ensuring that the training makes repeated use of the information presented, and providing for and encouraging elaboration of the material during acquisition as well as during the retention interval.

Other factors that have been shown to induce forgetting include interference by competing material that has been previously or subsequently acquired and the events encountered by individuals between learning and the retention test. Information acquired during this interval may impair retention, while simple rehearsal or reexposure may facilitate retention. Additional factors influencing skill and knowledge retention include the length of the retention interval, the methods used to assess retention, and individual difference variables among trainees. The absolute amount of skill/knowledge decay, for example, tends to increase with time, while the rate of forgetting declines over time. Researchers have also postulated a critical period for skills loss during a nonutilization period at between six months and one year after training (O'Hara 1990). Psychomotor flight skills, for example, are retained many months longer than procedural flight skills (Prophet 1976), and decay of flight skills begins to accelerate after a six-month nonutilization period (Ruffner et al. 1984).

### 3.6.3. *Transfer*

The topic of transfer-of-training is integrally related to other training issues, such as learning, memory, retention, cognitive processing, and conditioning; these fields make up a large subset of the subject matter of applied psychology (Swezey and Llaneras 1997). In general, the term *transfer-of-training* concerns the way in which previous learning affects new learning or performance. The central question is how previous learning transfers to a new situation. The effect of previous learning may function either to improve or to retard new learning. The first of these is generally referred to as positive transfer, the second as negative transfer. (If new learning is unaffected by prior learning, zero transfer is said to have occurred.) Many training programs are based upon the assumption that what is learned during training will transfer to new situations and settings, most notably the operational environment. Although U.S. industries are estimated to spend billions annually on training and development, only a fraction of these expenditures (not more than 10%) are thought to result in performance transfer to the actual job situation (Georgenson 1982). Researchers, therefore, have sought to determine fundamental conditions or variables that influence transfer-of-training and to develop comprehensive theories and models that integrate and unify knowledge about these variables.

Two major historical viewpoints on transfer exist. The identical elements theory (first proposed by Thorndike and Woodworth 1901) suggests that transfer occurs in situations where identical elements exist in both original and transfer situations. Thus, in a new situation, a learner presumably takes advantage of what the new situation offers that is in common with the learner's earlier experiences. Alternatively, the transfer-through-principles perspective suggests that a learner need not necessarily be aware of similar elements in a situation in order for transfer to occur. This position suggests that previously used "principles" may be applied to occasion transfer. A simple example involves the principles of aerodynamics, learned from kite flying by the Wright brothers, and the application of these principles to airplane construction. Such was the position espoused by Judd (1908), who suggested that what makes transfer possible is not the objective identities between two learning tasks, but the appropriate generalization in the new situation of principles learned in the old. Hendriksen and Schroeder (1941) demonstrated this transfer-of-principles philosophy in a series of studies related to the refraction of light. Two groups were given practice shooting at an underwater target until each was able to hit the target consistently. The depth of the target was then changed, and one group was taught the principles of refraction of light through water while the second was not. In a subsequent session of target shooting, the trained group performed significantly better than did the untrained group. Thus, it was suggested that it may be possible to design effective training environments without a great deal of concern about similarity to the transfer situation, as long as relevant underlying principles are utilized.

A model developed by Miller (1954) attempted to describe relationships among simulation fidelity and training value in terms of cost. Miller hypothesized that as the degree of fidelity in a simulation increased, the costs of the associated training would increase as well. At low levels of fidelity, very little transfer value can be gained with incremental increases in fidelity. However, at greater levels of fidelity, larger transfer gains can be made from small increments in fidelity. Thus, Miller hypothesized a point of diminishing returns, where gains in transfer value are outweighed by higher costs. According to this view, changes in the requirements of training should be accompanied by corresponding changes in the degree of fidelity in simulation if adequate transfer is to be provided. Although Miller did not specify the appropriate degree of simulation for various tasks, subsequent work (Alessi 1988) suggests that the type of task and the trainee's level of learning are two parameters that interact with Miller's hypothesized relationships. To optimize the relationship among fidelity, transfer, and cost, therefore, one must first identify the amount of fidelity of simulation required to obtain large amount of transfer and the point where additional increments of transfer are not worth the added costs.

Another model, developed by Kinkade and Wheaton (1972), distinguishes among three components of simulation fidelity: equipment fidelity, environmental fidelity, and psychological fidelity. Equipment fidelity refers to the degree to which a training device duplicates the "appearance and feel" of the operational equipment. This characteristic of simulators has also been termed physical fidelity. Environmental, or functional, fidelity refers to the degree to which a training device duplicates the sensory stimulation received from the task situation. Psychological fidelity (a phenomenon that Parsons [1980], has termed "verisimilitude") addresses the degree to which a trainee perceives the training device as duplicating the operational equipment (equipment fidelity) and the task situation (environmental fidelity). Kinkade and Wheaton postulated that the optimal relationship among levels of equipment, environmental, and psychological fidelity varies as a function of the stage of learning. Thus, different degrees of fidelity may be appropriate at different stages of training.

The relationship between degree of fidelity and amount of transfer is complex. Fidelity and transfer relationships have been shown to vary as a function of many factors, including instructor ability, instructional techniques, types of simulators, student time on trainers, and measurement techniques (Hays and Singer 1989). Nevertheless, training designers often favor physical fidelity in a training device, rather than the system's functional characteristics, assuming that increasing levels of physical fidelity are associated with higher levels of transfer. The presumption that similarity facilitates transfer can be traced to the identical elements theory of transfer first proposed by Thorndike and Woodworth (1901). In fact, numerous studies show that similarity does not need to be especially high in order to generate positive transfer-of-training (see Hays 1980; Provenmire and Roscoe 1971; Valverde 1973). Hay's (1980) review of fidelity research in military training found no evidence of learning deficits due to lowered fidelity, and others have shown an advantage for lower-fidelity training devices, suggesting that the conditions of simulation that maximize transfer are not necessarily those of maximum fidelity. One possible reason for this may be that widespread use of such well-known techniques as corrective feedback and practice to mastery in simulators and training devices, which may act to increase transfer, may also decrease similarity. A second possibility is that introducing highly similar (but unimportant) components into a simulation diverts a student's attention from the main topics being trained, thereby causing reduced transfer, as is the case with manipulating the fidelity of motion systems in certain types of flight simulators (Lintern 1987; Swezey 1989).

A recent review by Alessi (1988) devotes considerable attention to debunking the myth that high fidelity necessarily facilitates transfer. After reviewing the literature in the area, Alessi concluded that

transfer-of-training and fidelity are not linearly related but instead may follow an inverted-U-shaped function similar to that suggested by the Yerkes–Dodson law (see Swezey 1978; Welford 1968), which relates performance to stress. According to this view, increases in similarity first cause an increase and then a corresponding decrease in performance, presumably as a function of increasing cognitive load requirements associated with decreased stimulus similarity.

Current issues and problems in transfer span four general areas: (1) measurement of transfer, (2) variables influencing transfer, (3) models of transfer, and (4) application of our knowledge of transfer to applied settings. Research on transfer needs to be conducted with more relevant criterion measures for both generalization and skill maintenance. Hays and Singer (1989) provide a comprehensive review of this domain. A large proportion of the empirical research on transfer has concentrated on improving the design of training programs through the incorporation of learning principles, including identical elements, teaching of general principles, stimulus variability, and various conditions of practice. A critique of the transfer literature conducted by Baldwin and Ford (1988) also indicates that research in the area needs to take a more interactive perspective that attempts to develop and test frameworks to incorporate more complex interactions among training inputs. Many studies examining training design factors, for example, have relied primarily upon data collected using college students working on simple memory or motor tasks with immediate learning or retention as the criterion of interest. Generalizability, in many cases, is limited to short-term simple motor and memory tasks.

### 3.7. Training Teams and Groups

Because teams are so prevalent, it is important to understand how they function (Swezey and Salas 1992). Teamwork can take on various orientations, including behavioral and cognitive perspectives. Behaviorally oriented views of the teamwork process deal with aspects involved in coordinated events among members that lead to specific functions (e.g., tasks, missions, goals, or actions) that a team actually performs in order to accomplish its required output. Cognitive views of the teamwork domain deal with a team's shared processing characteristics, such as joint knowledge about a task, joint ability to perform aspects of the work, motivation levels, personalities, and thought processes (Salas et al. 1995). The extent to which mental models are shared among team members may help in understanding and training teamwork skills (Cannon-Bowers and Salas 1990). In order for members of a team to coordinate effectively with others, they must share knowledge of team requirements. This, in turn, can help to coordinate a team effectively. Since the range of each individual team member's abilities varies widely, the overall level of competency within a team is often dependent upon shared skills. An individual's ability to complete a team function competently and convey information to other members accurately is a strong influence on the overall team structure and performance.

Swezey et al. (1994) suggest that in order to complete a team task accurately, four things must happen: (1) there must be an exchange of competent information among team members; (2) there must be coordination within the team structure of the task requirements; (3) periodic adjustments must be made to respond to task demands; and (4) a known and agreed-upon organization must exist within the group.

It is important to consider such issues as team size, task characteristics, and feedback use when designing team training programs. It has been shown that team size has a direct influence on the performance and motivation of teams. As the size of the team increases, team resources also increase (Shaw 1976). Due to the increase in information available to larger teams, individual team members have a more readily accessible pool of information for use in addressing team tasks. This in turn can provide increases in creativity, the processing of information, and overall team effectiveness (Morgan and Lassiter 1992). However, the unique aspects associated with larger teams may do more harm than good due to the fact that the increase in size may also pose problems. Various studies (Indik 1965; Gerard et al. 1968) have found that problems in communication and level of participation may increase due to increases in team size. These two issues, when combined, may diminish team performance by placing increased stress on a team in the areas of coordination and communication workload (Shaw 1976).

Task goal characteristics are another important contributor to team performance. Given a task that varies in difficulty, it has been found that goals that appear to be challenging and that target a specific task tend to have a higher motivational impact than those that are more easily obtainable (Ilgen et al. 1987; Gladstein 1984; Goodman 1986; Steiner 1972). According to Ilgen et al. (1987), motivation directs how much time an individual is willing to put forth in accomplishing a team task.

The use of feedback is another factor that influences team motivation and performance. Feedback helps to motivate an individual concerning recently performed tasks (Salas et al. 1992). Researchers concur that the feedback should be given within a short period of time after the relevant performance (Dyer 1984; Nieva et al. 1978). The sequence and presentation of the feedback may also play significant roles in motivation. Salas et al. (1992) have noted that during the early stages of training,

feedback should concentrate on *one* aspect of performance; however, later in time it should concentrate on *several* training components. It has been found that sequencing helps team training in that teams adjust to incorporate their feedback into the task(s) at hand (Briggs and Johnston 1967).

According to Salas and Cannon-Bowers (1995), a variety of methods and approaches are currently under development for use in building effective team training programs. One such technique is the developing technology of team task analysis (Levine and Baker 1991), a technique that, as it matures, may greatly facilitate the development of effective team training. The technology provides a means to distinguish team learning objectives for effective performance from individual objectives and is seen as a potential aid in identifying the teamwork-oriented behaviors (i.e., cues, events, actions) necessary for developing effective team training programs.

A second area is team-performance measurement. To be effective, team-performance measurement must assess the effectiveness of various teamwork components (such as team-member interactions) as well as intrateam cognitive and knowledge activities (such as decision making and communication) in the context of assessing overall performance of team tasks. Hall et al. (1994) have suggested the need for integrating team performance outcomes into any teamwork-measurement system. As team feedback is provided, it is also necessary to estimate the extent to which an individual team member is capable of performing his or her specific tasks within the team. Thus, any competently developed team performance-measurement system must be capable of addressing both team capabilities and individual capabilities separately and in combination.

Teamwork simulation exercises are a third developing technology cited by Salas and Cannon-Bowers (1995). The intent of such simulations is to provide trainees with direct behavioral cues designed to trigger competent teamwork behaviors. Essential components of such simulations include detailed scenarios or exercises where specific teamwork learning objectives are operationalized and incorporated into training.

Salas and Cannon-Bowers (1995) have also identified three generic training *methods* for use with teams; information based, demonstration based, and practice based. The information-based method involves the presentation of facts and knowledge via the use of such standard delivery techniques as lectures, discussions, and overheads. Using such group-based delivery methods, one can deliver relevant information simultaneously to large numbers of individuals. The methods are easy to use and costs are usually low. Information-based methods may be employed in many areas, such as helping teammates understand what is expected of them, what to look for in specific situations, and how and when to exchange knowledge.

Demonstration-based methods are performed, rather than presented, as are information-based methods. They offer students an opportunity to observe behaviors of experienced team members and thus of the behaviors expected of them. Such methods help to provide shared mental models among team members, as well as examples of how one is expected to handle oneself within complex, dynamic, and multifaceted situations (Salas and Cannon-Bowers 1995).

Practice-based methods are implemented via participatory activities such as role playing, modeling, and simulation techniques. These methods provide opportunities to practice specific learning objectives and receive feedback information. With such methods, trainees can build upon previous practice attempts until achieving the desired level(s) of success.

A final category of developmental teamwork technologies, according to Salas and Cannon-Bowers (1995), involves teamwork training implementation strategies. These authors discuss two such strategies: cross-training and team coordination training. In their view, cross-training, training all team members to understand and perform each other's tasks, is an important strategy for integrating inexperienced members into experienced groups. Team coordination training addresses the issue that each team member has specific duties and that those duties, when performed together, are designed to provide a coordinated output. Team coordination training involves the use of specific task-oriented strategies to implement coordination activities and has been successfully applied in the aviation and medical fields.

### 3.8. Training Evaluation

Training evaluation may serve a variety of purposes, including improving training and assessing the benefits of training. Formative evaluation consists of processes conducted while training is being developed. Summative evaluation consists of processes conducted after training has been implemented.

Because training typically does not remain static (a job's tasks are often modified, replaced, or added), the distinction between formative evaluation and summative evaluation is sometimes blurred (see Goldstein 1993). Formative evaluation focuses on improving training prior to implementation. It may also serve the purpose of assessing benefits of training in a preliminary way. Summative evaluation focuses on assessing the benefits of training to determine how training outcomes improve on-job performance. It may also serve the purpose of improving future training iterations. In terms

of instructional systems design, summative evaluation completes the cycle that began with analysis. Figure 1 shows a simplified version of this cycle.

Both formative and summative aspects of evaluation feed back into the training analysis phase, which showed why training was necessary, the performances that training presumably would change, and the job outcomes that would be affected. Formative evaluation typically includes internal quality checks (such as review by subject matter experts, review against instructional standards, and editing for organization and clarity), as well as pilot testing. Formative evaluation may occur throughout the process shown in Figure 1. Summative evaluation includes assessments of after-training outcomes, both immediately following the conclusion of a training program and at later intervals.

Kirkpatrick (1994) has developed a widely used classification taxonomy of evaluation levels. Table 2 shows these levels and how they apply to both formative and summative evaluation.

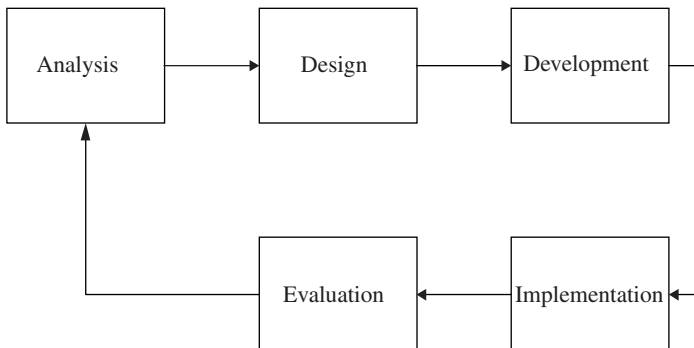
Note that pilot testing conducted during formative evaluation should address, at a minimum, levels 1 and 2 of Kirkpatrick's taxonomy. Summative evaluation should address all four levels. Kirkpatrick (1994) has suggested that all four levels must be addressed for evaluation to be useful to organizations. Although using the level 1 evaluation standard as a measure of satisfaction can be useful, using it as an indicator of training effectiveness is misleading. Clark (1982), for instance, showed that about one-third of the time, a negative correlation exists between a learner's end-of-course self-rating and measures of learning.

Level 2 evaluation is obviously beneficial for determining whether training produces the learning it is intended to produce. As discussed above, we recommend criterion-referenced testing as the appropriate source of level 2 data for training (as opposed to education) evaluation. End-of-course criterion-referenced tests are designed to show whether or not trainees can perform adequately on measures of each course objective. In criterion-referenced measurement, the operative question is, "Can trainees perform the required task or not?" Contrast this with norm-referenced testing, which addresses the question, "How do learners' levels of knowledge compare?" In practice, a norm-referenced test could show a percentage of students excelling, while a criterion-referenced test on the same information could show all students as failing to meet certain training standards.

Conducted with a criterion-referenced approach, level 2 evaluation shows what trainees can do at the end of a course, but not what they will do on the job. There are several reasons for this, as noted above. First, there is the issue of retention over time. If learners do not have the opportunity to practice what they have learned in training, retention will decrease. Consequently, for training to be effective, it should be conducted close to the time when trainees need to use what they have learned. Transfer issues, as previously described, may also affect actual job performance and should be carefully considered in training design.

Another reason that level 2 evaluation may not predict how trainees will perform later in time is that outside factors may inhibit application of what is learned. As noted previously, organizations may lack appropriate motivational systems and environmental support for specific behaviors. Thus, if trainees are taught to perform behavior  $y$  in response to stimulus  $x$  in training but are required to perform behavior  $z$  in response to that stimulus on the job, the behavior learned in training will soon be extinguished. Similarly, if employees are taught to use certain tools during training but those tools are absent from the work environment, the tool-related behaviors learned during training will not be used.

Kirkpatrick's level 3 evaluation seeks to measure changes in behavior after training. Because factors other than training may influence later behaviors, measuring these changes successfully may



**Figure 1** Evaluation's Role in an ISD Model.



**TABLE 2 Kirkpatrick’s Levels Applied to Formative and Summative Evaluation**

Level of Evaluation	Type of Evaluation	
	Formative	Summative
<b>1. Reaction:</b> Participants’ reaction to course (customer satisfaction).	“Dog food principle”: Even if it’s very nutritious, if the dogs don’t like it, they won’t eat it. Make sure participants are satisfied with the course.	Reactions can be used to distinguish between instructors, setting, audience, and other factors (e.g., are ratings for some instructors consistently higher than for others?).
<b>2. Learning:</b> Change measured at end of course.	Does the course meet its learning objectives? If it doesn’t, revise and repilot test until it does.	Use end-of-course measures to determine needs for remediation and for improving future iterations of course.
<b>3. Behavior:</b> Change measured on the job.	Typically not included in formative evaluation.	If the course doesn’t change behavior on the job, the investment in development has been wasted. Analyze causes for lack of change. Modify course, motivational factors, and/or environmental factors as necessary.
<b>4. Results:</b> Impact of change on job outcomes.	Typically not included in formative evaluation.	If the behavior change occurs but does not improve the job outcomes it targeted, the training may be flawed and/or external factors may be affecting the outcome. Reanalyze needs.

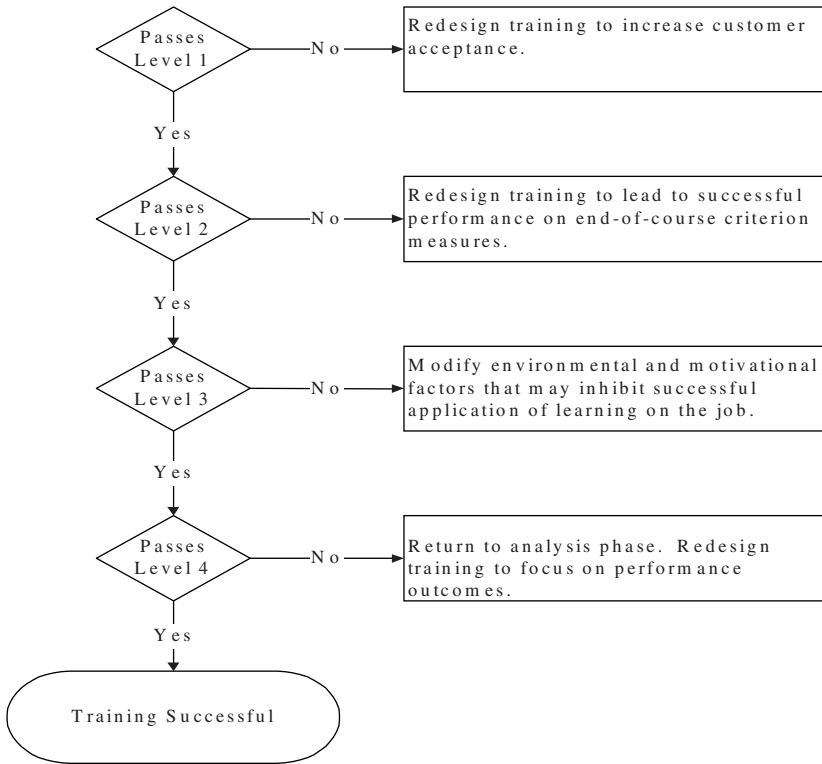
require an experiment. Table 3 shows an example of an experimental design permitting a level 3 evaluation.

Using the design shown in Table 3, training has been randomly assigned to group B. Baseline performance of both groups is then measured at the same time, before group B receives training. If the assignment is truly random, there should be no significant differences between the performance of trainees in groups A<sub>1</sub> and B<sub>1</sub>. At the same point after training, job performance of both groups is again measured. If all other factors are equal (as they should be in random assignment), differences between groups A<sub>2</sub> and B<sub>2</sub> will show the effect of training on job behavior. Depending on the type of measurement, one may use various statistical indices, such as a chi squared test or phi coefficient as a test of statistical significance (see Swezey 1981 for a discussion of this issue.)

This design does not use level 2 evaluation data. However, by modifying the design so that measures of A<sub>1</sub> and B<sub>1</sub> behaviors occur at the end of training (rather than prior to training) and measures of A<sub>2</sub> and B<sub>2</sub> occur later (e.g., at some interval after training), one can use level 2 evaluation data to assess the relative effect of nontraining factors on job performance. In this case (level 2 evaluation), B<sub>1</sub> performance should be significantly higher than A<sub>1</sub> performance. But B<sub>2</sub> performance may not be significantly higher than A<sub>2</sub> performance, depending upon whether nontraining factors inhibit job performance. Again, statistical comparisons are used to determine significance of the effect(s).

**TABLE 3 Experimental Design for Level 3 Evaluation**

Points of Measurement	Experimental Condition	
	No Training	Training
Before Training	A <sub>1</sub>	B <sub>1</sub>
After Training	A <sub>2</sub>	B <sub>2</sub>



**Figure 2** Remediating Failures at Each Level of Evaluation.

Kirkpatrick's level 4 evaluation addresses the basic question, "So what?" Even if evaluation at levels 1 through 3 show positive results, if training does not lead to desired results on job performance, it may not be the proper solution or worth the expense. For example, if training is designed to lead to performance that reduces error rates, and these error rates do not decline adequately, training may not have been the correct solution to the problem situation. Figure 2 provides a brief overview of how to address failure at each of Kirkpatrick's four levels of evaluation.

## 4. DEVELOPMENT

The distinction between training and development is ill defined. For purposes of this chapter, we define development in organizational contexts as "Plans and activities intended to achieve long-term changes in employees and their organizations." As such, development may include education, training, organizational realignment, values clarification, modifications to feedback and reward systems, mentoring, and a host of other interventions. In this section, we briefly mention a few forms of development in terms of research results.

### 4.1. Individual Performance Enhancement

In the 1940s and 1950s, much research on individual performance enhancement focused on personality development. By the 1960s, however, it had become clear that a wide variety of personalities could function successfully in nearly all jobs. One of the most widely used psychometric tools in business and industry, the Myers-Briggs Type Indicator (MBTI), purports to measure personality types. However, both Druckman and Bjork (1991) and Pittenger (1993) showed that data obtained via the MBTI have very little predictive validity in terms of job performance. Although research into the predictive value of personality testing continues today, it remains questionable whether development that focuses on personality traits as predictors will lead to significant results with respect to job performance.

Recently, the issue of competencies has also become popular in training circles. Most research on competencies can be divided into two types:

1. Those that address specific job skills, e.g., “can calibrate a variety of instruments used to measure tensile strength”
2. Those that address traits, (sometimes disguised in behavioral terms), e.g., “demonstrates the ability to work well with others in a variety of situations”

Although training may be an appropriate development activity for the first type of competency, the second may not benefit from *any* development activity. For development activities to be useful, their outcomes must be clearly specified. If “works well with others” is defined in operational terms (e.g., others report that the individual contributes well in team situations), then development activities, such as training, coaching, and mentoring, may be appropriate.

A variety of experiential forms of training are in use as development activities. These range from wilderness survival courses to games and specially designed on-the-job training experiences that focus on an individual’s performance. Spitzer (1986) reviews the preparation, conduct, and follow up of a variety of specially designed on-the-job training experiences. Most such experiential approaches have not been validated via careful research. Thiagarajan (1980) also summarizes advantages and disadvantages of games and other forms of experiential learning. He notes that, while experiential learning is often shown to produce the same results as more conventional forms of training, it occasionally leads to poorer skill acquisition. However, he argues that because such experiential techniques are often fun, they may gain trainee acceptance and therefore be a good vehicle for individual skill enhancement. We tend to disagree with this view.

Coaching and mentoring are other methods of performance enhancement used in development situations. In coaching, a coach works with a student during one or more sessions to help the student develop and apply skills on the job. Pearlstein and Pearlstein (1991) present an ISD-based approach to coaching, and Fournies (1987) describes a wide variety of contexts for coaching. Generally, mentoring is a long-term relationship between an experienced and a less-experienced individual that is designed to bring about changes in behavior via counseling, feedback, and a variety of other techniques. Murray (1991) provides an overview of factors and techniques contributing to effective mentoring.

Career planning as a development activity can also be an effective means of enhancing performance. In career planning, individuals and their supervisors and/or human resources departments develop and plan a series of steps to help align individual goals with organizational goals and career progression. These steps may include training and education (either on or off the job), coaching, mentoring, special assignments, and job rotations. Schein (1978) provide a detailed model for career planning. While most employee-assessment systems appear to be a sound basis for individual performance enhancement, their unsuccessful implementation may work against job performance. According to Scherkenbach (1988), Deming’s total quality management (TQM) process argues against such systems precisely because poor implementation can cause major roadblocks to success of both individuals and organizations. Since most employee assessment systems are linked to pay decisions, employees who admit to development needs may, in effect, be arguing against raises. This link thus creates both unsuccessful implementations and possible conflicts of interest.

## 4.2. Organizational Performance Enhancement

Two commonly practiced forms of organizational performance enhancement are management and leadership development and organization development. Management and leadership development focuses on improving organizational outcomes by improving the performance of those responsible for the outcomes. Organization development focuses on developing a total organization, as opposed to some particular aspect(s) of the organization, such as its leadership, management, or employee skills (see Gallessich 1982; Vaill 1971). Pearlstein (1991) argues that leadership development cannot be separated from organization development because leaders cannot be developed outside of the context of organization-wide goals for development.

### 4.2.1. Management and Leadership Development

Stodgill (1974), in a summary of over 3000 books and articles, found hundreds of definitions of leaders and managers. Some of these define management as focusing on support of organizational activities in their current state, as contrasted with leadership, which focuses on changing an organization to meet new challenges. In this context, a distinction is often made based upon topic areas, with management focusing on basics such as making assignments, delegating, and feedback and leadership focusing on more advanced areas such as communicating vision, establishing shared values, and individual employee empowerment. However, in the context of current organizational challenges, arbitrary distinctions between management and leadership are increasingly less relevant. Vaill (1989) asserts that organizations often find themselves, like kayakers, in “permanent white water.” This is even more the case now, with today’s greater impacts of increasing globalization and rapidly accelerating technology.

Pearlstein (1992) notes that “Research (on leadership and management) . . . does not show conclusively that any leadership development training approach based on skill-building has led to long-term improvements in organizational functioning. Research does not show conclusively that any leadership development approach based on personal/awareness development or values clarification has led to long-term organizational improvements.” Howard and Bray (1988) summarize the research of over 30 years of research on longitudinal management development studies. Complex and surprising results are the rule. For example, interpersonal skills, considered by many assessment centers to be one of the most critical predictors of management success, decline over time for many managers. Worse, managers selected in the 1980s had lower levels of interpersonal skills than those selected in the 1950s. Similarly, Kotter (1982) reports on intensive studies of 15 general managers and concludes that training does not seem as important to leaders’ development as do childhood and early work experiences. McCall et al. (1988) summarize the responses of nearly 400 corporate leaders on factors important to their development and note that leaders cited a 12:1 ratio of nontraining to training factors. Finally, Pearlstein (1991) indicates that management and leadership development programs are usually generic, offered out of the context of specific management and leadership jobs, do not last long enough, and focus either on skills or on personal development but not on both.

#### **4.2.2. Organization Development**

Organization development (OD) is practiced in many large organizations today. OD practitioners use a form of experimentation called action research, which includes five major phases:

1. Reaching agreement on goals and procedures of the OD activity
2. Designing ways to diagnose needs for organizational change
3. Presenting diagnostic findings in a useful way
4. Using the diagnostic findings to plan an organizational change strategy.
5. Applying the change strategy and evaluating its outcomes

Both Lippitt et al. (1958) and Argyris (1970) provide examples of using this phased approach. The examples have in common behavioral approaches to collecting the data on which to base the change strategies. Pearlstein (1997) provides nine OD principles for those who wish to practice within a human performance technology context:

1. Develop work units (not necessarily the entire organization).
2. Make sure that you are interacting with the appropriate decision maker.
3. Start by reaching agreement on desired outcomes, measurements, roles, and activities.
4. Insist on behavioral measurement, but first listen carefully to the client’s statements of values and visions.
5. Make sure that the client considers the impacts of change on individuals, work units, the entire organization, and the larger community of which the organization is a part.
6. Meet with all key players, both individually and in groups.
7. Seek consensus but don’t withhold expertise.
8. Give interventions sufficient time to work.
9. Close well with the client by helping to: (a) collect and disseminate summative evaluation data, (b) celebrate organizational successes, and (c) plan next steps.

## **5. THE FUTURE OF SELECTION, TRAINING, AND DEVELOPMENT**

We do not claim to hold a crystal ball; however, some aspects of selection, training, and development are likely to stay the same, while others are likely to change. In our view, those that will stay the same are based on well-established research results and apply independently of increasing globalization and accelerating technological development. Those that will change lack these attributes.

### **5.1. Selection**

Selection tools will likely remain focused on measuring past and current behavior. Some tools that rely on assessing current performance may be slow to change because of difficulties associated with simulating tasks remotely. Many tools will become more widely used over the Internet. That is, organizations will increasingly use the Internet as a means for obtaining information previously gained via application forms, paper-based tests, interviews, and other selection devices. For example, many organizations now use Internet-based application forms that feed candidates’ responses in standard formats directly into databases, which are then mined to match candidates’ qualifications with job requirements. Similarly, because of globalization, it is likely that selection tools such as application

forms, tests, and interviews will also become increasingly available in multilingual and multicultural versions. Further, simulations and work sample assessments will be increasingly used, also in different culturally oriented and language-based forms.

## 5.2. Training

The basic development process for training is likely to remain some form of systems-oriented instructional model, although cognitively oriented technologies will increasingly be developed. The well-researched ISD approach has been adapted for use in many cultures and leads to reliable outcomes regardless of the specific type of technology employed. As cognitive approaches become better able to predict factors that build expertise, they will be more frequently used in task analysis. Results may be used to design training that is more useful for “far transfer” situations—that is, to tasks relatively dissimilar to those used as examples during training (see Clark 1992). Still, results of cognitive analysis are likely to be used within the context of a systems-oriented instructional model.

Training delivery and administration systems are very likely candidates for change. Computer-based training (CBT), now widely available in CD-ROM format, is increasingly more available in other disk-based formats (e.g., DVD) and on the Internet as Web-based training (WBT). Limitations on processing power, memory, storage, and bandwidth are all decreasing rapidly. This is leading to an increased ability for CBT, regardless of its mode of delivery, to offer high levels of simulation. Streaming audio and streaming video, voice recognition, and other technological advances are rapidly improving simulation fidelity. Similarly, advances in manipulanda (e.g., tactile gloves, aircraft yokes) are also increasing simulation fidelity. The result is that training delivered to remote sites can use both simulation and work sample presentation more effectively.

Other distance learning approaches, such as virtual classrooms, are being used increasingly. Both synchronous (real-time classroom) and asynchronous (classes via forums and message boards) types of distance learning will become widely used. For example, it is likely that organizations will increasingly incorporate university courses directly into work environments. This can be done globally with, for example, Saudi managers taking economics courses offered by Japanese universities.

There are negative aspects to the rapid spread of distance learning techniques. Often media values are emphasized at the expense of sound instructional practice. For example, much CBT available today does not make use of the technology for branching to different points in instruction based upon learner responses. Although the technology for branching has existed since the 1960s (see, e.g., Markle 1969), much CBT available today is linear. Although some argue that learner control (the ability to move through lessons at will) is more important than programmed control, research findings suggest otherwise. Clark (1989), for instance, showed that given a choice of approach (structured or unstructured), CBT users were more likely to choose the alternative least useful for skill acquisition. Novice learners, who may have benefited from a structured approach, often chose an unstructured one, while experts, who could use their advanced ability to explore advantageously, were more likely to choose structured approaches. Similarly, the advance of virtual classrooms in corporate settings may lead to an emphasis on education, as opposed to training. Thus, employees may learn more generalities about topics, at the expense of learning how to perform tasks.

One recent development in training has been slow to evolve. Electronic performance support systems (EPSS), which combine online help, granular training, and expert advice functions, were popularized by Gery (1991). Although used in some corporations today, EPSS use does not seem to be accelerating, possibly because development costs are relatively high. In an EPSS, help is directly available on employee computers. In addition to highly contextual online help, brief, focused training modules are offered when users repeatedly request help on the same topics. Such systems use artificial intelligence techniques, such as algorithms based upon data derived from users' experiences, to offer expert advice. Thus, a user who wanted to advise a consumer on what lumber to purchase for a deck could select an expert advisor function to input the consumer's space requirements and house dimensions, and the computer could immediately provide appropriate sketches and lumber requirements. Such systems can dramatically reduce amounts of time spent in formal training.

## 5.3. Development

Like training, approaches for individual, management, and leadership development are likely to be offered increasingly on the Internet via virtual classrooms. Since many organizations continue to spend a large amount on development activities, there is little reason to suppose that they will discontinue such expenditures.

### Acknowledgments

Portions of this chapter were adapted from Swezey and Llaneras (1997). The authors would like to thank Gloria Pearlstein and Lisa Dyson for their help with the preparation of this chapter.

## REFERENCES

- Alessi, S. M. (1988), "Fidelity in the Design of Instructional Simulations," *Journal of Computer-Based Instruction*, Vol. 15, No. 2, pp. 40-47.
- Allen, W. H. (1973), "Research in Educational Media," in *Educational Media Yearbook*, J. W. Brown, Ed., R.R. Bowker, New York.
- Alluisi, E. A. (1991), "The Development of Technology for Collective Training: SIMNET, a Case History," *Human Factors*, Vol. 33, No. 3, 343-362.
- American Educational Research Association (AERA), American Psychological Association (APA), and National Council on Measurement in Education (NCME) (1999), *Standards for Educational and Psychological Testing*, American Psychological Association, Washington, DC.
- American Educational Research Association (1999), *Standards for Educational and Psychological Testing*, American Psychological Association, Washington, DC.
- Anderson, J. R. (1982), "Acquisition of Cognitive Skill," *Psychological Review*, Vol. 89, pp. 369-406.
- Anderson, J. R. (1990), *Cognitive Psychology and Its Implications*, 3rd Ed., W.H. Freeman, New York.
- Andrews, D. H., Waag, W. L., and Bell, H. H. (1992), "Training Technologies Applied to Team Training: Military Examples," in *Teams: Their Training and Performance*, R. W. Swezey and E. Salas, Eds., Ablex, Norwood, NJ.
- Argyris, C. (1970), *Intervention Theory and Method: A Behavioral Science View*, Addison-Wesley, Reading, MA.
- Bahrick, H. P. (1984), "Semantic Memory Content in Permastore: Fifty Years of Memory for Spanish Learned in School," *Journal of Experimental Psychology: General*, Vol. 113, pp. 1-29.
- Baldwin, T. T., and Ford, J. K. (1988), "Transfer-of-Training: A Review and Directions for Future Research," *Personal Psychology*, Vol. 41, pp. 63-105.
- Battig, W. F. (1966), "Facilitation and Interference," in *Acquisition of Skill*, E. A. Bilodeau, Ed., Academic Press, New York.
- Bloom, K. C., and Shuell, T. J. (1981), "Effects of Massed and Distributed Practice on the Learning and Retention of Second-Language Vocabulary," *Journal of Educational Research*, Vol. 74, No. 4, pp. 245-248.
- Bouzd, N., and Crawshaw, C. M. (1987), "Masses Versus Distributed Word Processor Training," *Applied Ergonomics*, Vol. 18, pp. 220-222.
- Branson, R. K., Rayner, G. T., Cox, J. L., Furman, J. P., King, F. J., and Hannum, W. J. (1975), "Interservice procedures for instructional systems development: Executive summary and model," ADA019486, U.S. Army Combat Arms Training Board, Fort Benning, GA.
- Briggs, G. E., and Johnston, W. A. (1967), "Team Training," NAVTRADEVCEEN-1327-1, AD-608 309, Naval Training Device Center, Port Washington, NY.
- Campbell, J. (1990), "An Overview of the Army Selection and Classification Project (Project A)," *Personnel Psychology*, Vol. 43, pp. 231-239.
- Cannon-Bowers, J. A., and Salas, E. (1990), "Cognitive Psychology and Team Training: Shared Mental Models in Complex Systems," Paper presented at the 5th Annual Conference of the Society for Industrial and Organizational Psychology, K. (Miami).
- Carlisle, K. E. (1986), *Analyzing Jobs and Tasks*. Educational Technology Publications, Englewood Cliffs, NJ.
- Clark, R. E. (1982), "Antagonism between Achievement and Enjoyment in ATI Studies," *Educational Psychology*, Vol. 17, No. 2, pp. 92-101.
- Clark, R. E. (1983), "Reconsidering Research on Learning from Media," *Review of Educational Research*, Vol. 53, pp. 445-459.
- Clark, R. E. (1989), "When Teaching Kills Learning: Research on Mathematics," in *Learning and Instruction: European Research in an International Context*, Mandl, H. N., Bennett, N., de Corte, E., and Freidrich, H. F., Eds., Vol. 2, Pergamon Press, London.
- Clark, R. E. (1992), "How the Cognitive Sciences are Shaping the Profession," in *Handbook of Human Performance Technology*, Stolovitch, H. D., and Keeps, E. J., Eds., Jossey-Bass, San Francisco.
- Clark, R. E. (1994), "Media Will Never Influence Learning," *Educational Technology Research and Development*, Vol. 42, No. 3, pp. 21-29.

- Cofer, C. N., Bruce, D. R., and Reicher, G. M. (1966), "Clustering in Free Recall as a Function of Certain Methodological Variations," *Journal of Experimental Psychology*, Vol. 71, pp. 858–866.
- Cohen, P. A., Ebeling, B. J., and Kulik, J. A. (1981), "A Meta-analysis of Outcome Studies of Visual-Based Instruction," *Educational Communication and Technology Journal*, Vol. 29, pp. 26–36.
- Cormier, S. M. (1984), "Transfer-of-training: An interpretive review," Tech. Rep. No. 608, Army Research Institute, Alexandria, VA.
- Craik, F. I. M., and Lockhart, R. S. (1972), "Levels of Processing: A Framework for Memory Research," *Journal of Verbal Learning and Verbal Behavior*, Vol. 11, pp. 671–684.
- Druckman, D., and Bjork, R. A., Eds. (1991), *In the Mind's Eye: Enhancing Human Performance*, National Academy Press, Washington, DC.
- Dwyer, D., Hall, J., Volpe, C., Cannon-Bowers, J. A., and Salas, E. (1992), "A Performance Assessment Task for Examining Tactical Decision Making under Stress," Tech. Rep. No. 92-002, U.S. Naval Training Systems, Center, Orlando, FL.
- Dyer, J. L. (1984), "Review on Team Training and Performance: A State-of-the-Art Review," in *Human factors review*, F. A. Muckler, Ed., The Human Factors Society, Santa Monica, CA.
- Ebbinghaus, H. (1913), *Memory: A Contribution to Experimental Psychology*, H. A. Ruger and C. E. Bussenius, Trans., Columbia University, New York (original work published 1885).
- Farr, M. J. (1987), *The Long-Term Retention of Knowledge and Skills: A Cognitive and Instructional Perspective*, Springer-Verlag, New York.
- Fletcher, J. D. (1990), "Effectiveness and Cost of Interactive Videodisc Instruction in Defense Training and Education," IDA Paper P-2372, Institute for Defense Analyses, Alexandria, VA.
- Fournies, F. F. (1987), *Coaching for Improved Work Performance*, Liberty Hall Press, Blue Ridge Summit, PA.
- Fraser, S. L., and Kroeck, K. G. (1989), "The Impact of Drug Screening on Selection Decisions," *Journal of Business Psychology*, Vol. 3, pp. 403–411.
- Gagné, E. D. (1978), "Long-Term Retention of Information Following Learning from Prose," *Review of Educational Research*, Vol. 48, pp. 629–665.
- Gallessich, J. (1982), *The Profession and Practice of Consultation*, Jossey-Bass, San Francisco.
- Georgenson, D. L. (1982), "The Problem of Transfer Calls for Partnership," *Training and Development Journal*, Vol. 36, No. 10, pp. 75–78.
- Gerard, H. B., Wilhelmy, R. A., and Conolley, E. S. (1968), "Conformity and Group Size," *Journal of Personality and Social Psychology*, Vol. 8, pp. 79–82.
- Gery, G. J. (1991), *Electronic Performance Support Systems: How and Why to Remake the Workplace through the Strategic Application of Technology*, Weingarten, Boston.
- Gilbert, T. F. (1978), *Human Competence: Engineering Worthy Performance*, McGraw-Hill, New York.
- Gladstein, D. L. (1984), "Groups in Context: A Model of Task Group Effectiveness," *Administrative Science Quarterly*, Vol. 29, pp. 499–517.
- Glaser, R. B. (1963), "Instructional Technology and the Measurement of Learning Outcomes: Some Questions," *American Psychologist*, Vol. 18, pp. 519–521.
- Goldstein, I. L. (1993), *Training in Organizations*, 3rd Ed., Brooks/Cole, Pacific Grove, CA.
- Goodman, P. S., Ed. (1986), *Designing Effective Work Groups*, Jossey-Bass, San Francisco.
- Guion, R. M. (1991), "Personnel Assessment, Selection, and Placement," in *Handbook of Industrial and Organizational Psychology*, M. D. Dunnette and L. M. Hough, Eds., Consulting Psychologists Press, Palo Alto, CA, pp. 327–397.
- Hall, J. K., Dwyer, D. J., Cannon-Bowers, J. A., Salas, E., and Volpe, C. E. (1994), "Toward Assessing Team Tactical Decision Making under Stress: The Development of a Methodology for Structuring Team Training Scenarios," in *Proceedings of the 15th Annual Interservice/Industry Training Systems and Education Conference* (Washington, DC), pp. 87–98.
- Hammer, E. G., and Kleiman, L. A. (1988), "Getting to Know You," *Personnel Administration*, Vol. 34, pp. 86–92.
- Harless, J. H. (1975), *An Ounce of Analysis (Is Worth a Pound of Objectives)*, Harless Performance Guild, Newman, GA.
- Hays, R. T. (1980), "Simulator Fidelity: A Concept Paper," Tech. Rep. No. 490, Army Research Institute, Alexandria, VA.
- Hays, R. T., and Singer, M. J. (1989), *Simulation Fidelity in Training System Design*, Springer, New York.

- Hendriksen, G., and Schroeder, W. H. (1941). "Transfer-of-Training in Learning to Hit a Submerged Target," *Journal of Educational Psychology*, Vol. 32, pp. 205–213 (cited in Hilgard 1962).
- Herzberg, F. (1974). "The Wise Old Turk," *Harvard Business Review*, Vol. 52, September–October, pp. 70–80.
- Hilgard, E. R. (1962), *Introduction to Psychology*, 3d Ed., Harcourt Brace & World.
- Hoban, C. F., and Van Ormer, E. B. (1950). "Instructional Film Research, 1918–1950," Tech. Rep. No. SDC 269-7-19, U.S. Naval Special Devices Center, Port Washington, NY (ERIC Document Reproduction Service No. ED 647 255).
- Hollenbeck, J. R., Segó, D. J., Ilgen, D. R., and Major, D. A. (1991). "Team Interactive Decision Exercise for Teams Incorporating Distributed Expertise (TIDE<sup>2</sup>): A Program and Paradigm for Team Research," Tech. Rep. No. 91-1, Michigan State University, East Lansing, MI.
- Howard, A., and Bray, D. (1988), *Managerial Lives in Transition: Advancing Age and Changing Times*, Guilford Press, New York.
- Hunter, J. E. (1986). "Cognitive Ability, Cognitive Aptitudes, Job Knowledge, and Job Performance," *Journal of Vocational Behavior*, Vol. 29, pp. 340–62.
- Hurlock, R. E., and Montague, W. E. (1982). "Skill Retention and Its Primary Implications for Navy Tasks: An Analytical Review," NPRDC SR 82-21, Navy Personnel Research & Development Center, San Diego.
- Ilgen, D. R., Shapiro, J., Salas, E., and Weiss, H. (1987). "Functions of Group Goals: Possible Generalizations from Individuals to Groups," Tech. Rep. No. 87-022, Naval Training Systems Center, Orlando, FL.
- Indik, B. P. (1965). "Organizational Size and Member Participation: Some Empirical Test of Alternatives," *Human Relations*, Vol. 18, pp. 339–350.
- Jacobs, J. W., Prince, C., Hays, R. T., and Salas, E. (1990). "A Meta-analysis of the Flight Simulator Training Research," NAVTRASYS-CEN TR-89-006, Naval Training Systems Center, Orlando, FL.
- Jarrard, S. W., and Wogalter, M. S. (1992). "Recognition of Non-studies Visual Depictions of Aircraft: Improvement by Distributed Presentation," in *Proceedings of the Human Factors Society 36th Annual Meeting* (Atlanta, October 12–16), pp. 1316–1320.
- Jensen, A. R. (1986). "G: Artifact or Reality?" *Journal of Vocational Behavior*, Vol. 29, pp. 301–331.
- Judd, C. H. (1908). "The Relation of Special Training and General Intelligence," *Educational Review*, Vol. 36, pp. 28–42.
- Kennedy, R. S., Jones, M. B., and Harbeson, M. M. (1980). "Assessing Productivity and Well-Being in Navy Work-Places," in *Proceedings of the 13th Annual Meeting of the Human Factors Association of Canada* (Point Ideal, Lake of Bays, ON), Human Factors Association of Canada, Downsview, ON, pp. 8–13.
- Kieras, D. E., and Bovair, S. (1984). "The Role of a Mental Model in Learning to Operate a Device," *Cognitive Science*, Vol. 8, pp. 255–273.
- Kinkade, R. G., and Wheaton, G. R. (1972). "Training Device Design," in *Human Engineering Guide to Equipment Design*, H. P. Van Cott and R. G. Kinkade, Eds., U.S. Printing Office, Washington, DC.
- Kirkpatrick, D. L. (1994), *Evaluating Training Programs: The Four Levels*. Barrett-Koehler, San Francisco.
- Kleiman, L. S., and Faley, R. H. (1990). "A Comparative Analysis of the Empirical Validity of Past and Present-Oriented Biographical Items," *Journal of Business Psychology*, Vol. 4, pp. 431–437.
- Kotter, J. (1982), *The General Managers*, Free Press, New York.
- Kulik, J. A., and Kulik, C. (1987). "Review of Recent Research Literature on Computer-Based Instruction," *Contemporary Educational Psychology*, Vol. 12, pp. 222–230.
- Kulik, J. A., Kulik, C., and Cohen, P. A. (1980). "Effectiveness of Computer-Based College Teaching: A Meta-analysis of Comparative Studies," *Review of Educational Research*, Vol. 50, pp. 525–544.
- Kyllonen, P. C., and Alluisi, E. A. (1987). "Learning and Forgetting Facts and Skills," in *Handbook of Human Factors*, G. Salvendy, Ed., John Wiley & Sons, New York.
- Larkin, J. (1981). "Enriching Formal Knowledge: A Model for Learning to Solve Textbook Physics Problems," in *Cognitive Skills and Their Acquisition*, J. R. Anderson, Ed., Erlbaum, Hillsdale, NJ.
- Lawler, E. E., III, Hackman, J. R., and Kaufman, S. (1973). "Effects of Job Redesign: A Field Experiment," *Journal of Applied Social Psychology*, Vol. 3, No. 1, pp. 49–62.



- Levine, E. L., and Baker, C. V. (1991), "Team Task Analysis: A Procedural Guide and Test of the Methodology," in *Methods and Tools for Understanding Teamwork: Research with Practical Implications*, E. Salas, Chair, Symposium presented at the 6th Annual Conference of the Society for Industrial and Organizational Psychology, St. Louis, MO.
- Lintern, G. (1987), "Flight Simulation Motion Systems Revisited," *Human Factors Society Bulletin*, Vol. 30, No. 12, pp. 1-3.
- Lippitt, R., Watson, J., and Westley, B. (1958), *The Dynamics of Planned Change*, Harcourt, Brace & World, New York.
- Loftus, G. R. (1985), "Evaluating Forgetting Curves," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 11, pp. 397-406.
- Luh, C. W. (1922), "The Conditions of Retention," *Psychological Monographs*, Whole No. 142, Vol. 31, No. 3.
- Mager, R. F. (1972), *Goal Analysis*, Lear Siegler/Fearon, Belmont, CA.
- Mager, R. F., and Pipe, P. (1970), *Analyzing Performance Problems or You Really Oughta Wanna*, Fearon, Belmont, CA.
- Markle, S. M. (1969), *Good Frames and Bad: A Grammar of Frame Writing*, 2nd Ed., John Wiley & Sons, New York.
- Martin, C. L., and Nagao, D. H. (1989), "Some Effects of Computerized Interviewing on Job Applicant Responses," *Journal of Applied Psychology*, Vol. 75, pp. 72-80.
- McCall, M. W., Lombardo, M. M., and Morrison, A. M. (1988), *Lessons of Experience: How Successful Executives Develop on the Job*, Lexington Books, Lexington, MA.
- McGeoch, J. A., and Irion, A. L. (1952), *The Psychology of Human Learning*, 2nd Ed., Longmans, Green & Co., New York.
- McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. L., and Ashworth, S. (1990), "Project A Validity Results: The Relationship between Predictor and Criterion Domains," *Personnel Psychology*, Vol. 43, pp. 335-354.
- Miller, R. B. (1954), *Psychological Considerations for the Design of Training Equipment*, American Institutes for Research, Pittsburgh.
- Morgan, B. B., Jr., and Lassiter, D. L. (1992), "Team Composition and Staffing," in *Teams: Their Training and Performance*, R. W. Swezey and E. Salas, Eds., Ablex, Norwood, NJ.
- Morris, N. M., and Rouse, W. B. (1985), "Review and Evaluation of Empirical Research in Troubleshooting," *Human Factors*, Vol. 27, No. 5, pp. 503-530.
- Mourier, P. (1998), "How to Implement Organizational Change That Produces Results," *Performance Improvement*, Vol. 37, No. 7, pp. 19-28.
- Murphy, K. R., Thornton, G. C., and Reynolds, D. H. (1990), "College Students' Attitudes toward Employee Drug Testing Programs," *Personnel Psychology*, Vol. 43, pp. 615-31.
- Murray, M. (1991), *Beyond the Myths and Magic of Mentoring: How to Facilitate an Effective Mentoring Program*, Jossey-Bass, San Francisco.
- Nieva, V. F., Fleishman, E. A., and Rieck, A. (1978), "Team Dimensions: Their Identity, Their Measurement, and Their Relationships," Contract No. DAHC19-78-C-0001, Response Analysis Corporation, Washington, DC.
- O'Hara, J. M. (1990), "The Retention of Skills Acquired through Simulator-Based Training," *Ergonomics*, Vol. 33, No. 9, pp. 1143-1153.
- Parsons, H. M. (1980), *Aspects of a Research Program for Improving Training and Performance of Navy Teams*, Human Resources Research Organization, Alexandria, VA.
- Parsons, P. J., Fogan, T., and Spear, N. E. (1973), "Short-Term Retention of Habituation in the Rat: A Developmental Study from Infancy to Old Age," *Journal of Comparative Psychological Psychology*, Vol. 84, pp. 545-553.
- Pearlstein, G. B., and Pearlstein, R. B. (1991), "Helping Individuals Build Skills through ISD-Based Coaching," Paper presented at Association for Education and Communication Technology Annual Conference (Orlando, FL, March).
- Pearlstein, R. B. (1991), "Who Empowers Leaders?" *Performance Improvement Quarterly*, Vol. 4, No. 4, pp. 12-20.
- Pearlstein, R. B. (1992), "Leadership Basics: Behaviors and Beliefs," Paper presented at the 30th Annual Conference of the National Society for Performance and Instruction (Miami).
- Pearlstein, R. B. (1997), "Organizational Development for Human Performance Technologists," in *The Guidebook for Performance Improvement: Working with Individuals and Organizations*, R. Kaufman, S. Thiagarajan, and P. MacGillis, Eds., Pfeiffer, San Francisco.

- Peterson, R. O., and Duffany, B. H. (1975), "Job Enrichment and Redesign," in *The Training and Development Handbook by AT&T*. AT&T, New York.
- Pittenger, D. J. (1993), "The Utility of the Myers-Briggs Type Indicator," *Review of Educational Research*, Vol. 63, No. 4, pp. 467-488.
- Prediger, D. J. (1989), "Ability Differences across Occupations: More than g," *Journal of Vocational Behavior*, Vol. 34, pp. 1-27.
- Prophet, W. W. (1976), "Long-Term Retention of Flying Skills: A Review of the Literature," HumRRO Final Report 76-35, ADA036977, Human Resources Research Organization, Alexandria, VA.
- Provenmire, H. K., and Roscoe, S. N. (1971), "An Evaluation of Ground-Based Flight Trainers in Routine Primary Flight Training," *Human Factors*, Vol. 13, No. 2, pp. 109-116.
- Rasmussen, J. (1979), "On the Structure of Knowledge: A Morphology of Mental Models in Man-Machine Systems Context," Report M-2192, Riso National Laboratory, Denmark.
- Rea, C. P., and Modigliani, V. (1988), "Educational Implications of the Spacing Effect," in *Practical Aspects of Memory: Current Research and Issues*, Vol. 1, M. M. Gruenberg, P. E. Morris, and R. N. Sykes, Eds., John Wiley & Sons, Chichester, pp. 402-406.
- Rossett, A. (1987), *Training Needs Assessment*, Educational Technology Publications, Englewood Cliffs, NJ.
- Ruffner, J., Wick, W., and Bickley, W. (1984), "Retention of Helicopter Flight Skills: Is There a Critical Period for Proficiency Loss?" in *Proceedings of the Human Factors Society 28th Annual Meeting*, Human Factors Society, Santa Monica, CA.
- Rumelhart, D. E., and Norman, D. A. (1978), "Accretion, Tuning and Restructuring: Three Modes of Learning," in *Semantic Factors in Cognition*, J. W. Cotton and R. L. Klatzky, Eds., Erlbaum, Hillsdale, NJ.
- Rummler, G. A., and Brache, A. P. (1990), *Improving Performance: How to Manage the White Space on the Organization Chart*, Jossey-Bass, San Francisco.
- Rummler, G. A., and Wilkins, C. L. (1999), "Performance Logic: Breaking the Code to Organization Performance," in *Proceedings of the Annual Conference of the International Association for Performance Improvement* (March), ISPI, Washington, DC.
- Russell, C. J., Mattson, J., Devlin, S. E., and Atwater, D. (1990), "Predictive Validity of Biodata Items Generated from Retrospective Life Experience Essays," *Journal of Applied Psychology*, Vol. 75, pp. 569-580.
- Sackett, P. R., and Decker, P. J. (1979), "Detection of Deception in the Employment Context: A Review and Critical Analysis," *Personnel Psychology*, Vol. 32, pp. 487-506.
- Salas, E., and Cannon-Bowers, J. A. (1995), "Methods, Tools, and Strategies for Team Training," in *Training for 21st Century Technology: Applications of Psychological Research*, M. A. Quinones and A. Dutta, Eds., APA Press, Washington, DC.
- Salas, E., Dickinson, T. L., Converse, S. A., and Tannenbaum, S. I. (1992), "Toward an Understanding of Team Performance and Training," in *Teams: Their Training and Performance*, R. W. Swezey and E. Salas, Eds., Ablex, Norwood, NJ.
- Salas, E., Cannon-Bowers, J. A., and Blickensderfer, E. L. (1995), "Team Performance and Training Research: Emerging Principles," *Journal of the Washington Academy of Sciences*, Vol. 83, No. 2, pp. 81-106.
- Salmoni, A. W., Schmidt, R. A., and Walter, C. B. (1984), "Knowledge of Results and Motor Learning: A Review and Critical Appraisal," *Psychological Bulletin*, Vol. 95, pp. 355-386.
- Schein, E. H. (1978), *Career Dynamics: Matching Individual and Organizational Needs*, Addison-Wesley, Reading, MA.
- Scherkenbach, W. W. (1988), *The Deming Route to Quality and Productivity: Road Maps and Roadblocks*, Mercury Press, Rockville, MD.
- Schmidt, R. A., and Bjork, R. A. (1992), "New Conceptualizations of Practice: Common Principles in Three Paradigms Suggest New Concepts for Training," *Psychological Science*, Vol. 3, No. 4, pp. 207-217.
- Schmidt, F., Ones, D., and Hunter, J. (1992), "Personnel Selection," *Annual Review of Psychology*, Vol. 43, pp. 627-670.
- Schneider, V. I., Healy, A. F., Ericsson, K. A., and Bourne, L. E., Jr. (1995), "The Effects of Contextual Interference on the Acquisition and Retention of Logical Rules," in *Learning and Memory of Knowledge and Skills*, A. F. Healy and L. E. Bourne, Jr., Eds., Sage, Newbury Park, CA.
- Schramm, W. (1977), *Big Media, Little Media*, Sage, Beverly Hills, CA.

- Shaw, M. E. (1976), *Group Dynamics: The Psychology of Small Group Behavior*, McGraw-Hill, New York.
- Shea, J. F., and Morgan, R. L. (1979), "Contextual Interference Effects on the Acquisition, Retention and Transfer of a Motor Skill," *Journal of Experimental Psychology: Human Learning and Memory*, Vol. 5, pp. 179–187.
- Spitzer, D. R. (1986), *Improving Individual Performance*, Educational Technology Publications, Englewood Cliffs, NJ.
- Steiner, I. D. (1972), *Group Process and Productivity*, Academic Press, New York.
- Stodgill, R. M. (1974), *Handbook of Leadership*, Free Press, New York.
- Streufert, S., and Swezey, R.W. (1985), "Simulation and Related Research Methods in Environmental Psychology," in *Advances in Environmental Psychology*, A. Baum and J. Singer, Eds., Erlbaum, Hillsdale, NJ, pp. 99–118.
- Swezey, R. W. (1978), "Retention of Printed Materials and the Yerkes-Dodson Law," *Human Factors Society Bulletin*, Vol. 21, pp. 8–10.
- Swezey, R. W. (1981), *Individual Performance Assessment: An Approach to Criterion-Referenced Test Development*, Reston, Reston, VA.
- Swezey, R. W. (1989), "Generalization, Fidelity and Transfer-of-Training," *Human Factors Society Bulletin*, Vol. 32, No. 6, pp. 4–5.
- Swezey, R. W., Hutcheson, T. D., Rohrer, M. W., Swezey, L. L., and Tirre, W. C. (1999), "Development of a Team Performance Assessment Device (TPAD): Final Report," InterScience America, Leesburg, VA (Report prepared under Contract No. F41624-97-C-5006 with the U.S. Air Force Research Laboratory, Brooks AFB, TX).
- Swezey, R. W., and Llaneras, R. E. (1992), "Validation of an Aid for Selection of Instructional Media and Strategies," *Perceptual and Motor Skills*, Vol. 74, p. 35.
- Swezey, R.W., and Llaneras, R.E. (1997), "Models in Training and Instruction," in *Handbook of Human Factors and Ergonomics*, 2nd Ed., G. Salvendy, Ed., John Wiley & Sons, New York. pp. 514–577.
- Swezey, R. W., and Salas, E. (1992), "Guidelines for Use in Team Training Development," in *Teams, Their Training and Performance*, R. W. Swezey and E. Salas, Eds., Ablex, Norwood, NJ.
- Swezey, R. W., Perez, R. S., and Allen, J. A. (1991), "Effects of Instructional Strategy and Motion Presentation Conditions on the Acquisition and Transfer of Electromechanical Troubleshooting Skill," *Human Factors*, Vol. 33, No. 3, pp. 309–323.
- Swezey, R. W., Meltzer, A. L., and Salas, E. (1994), "Issues Involved in Motivating Teams," in *Motivation: Theory and Research*, H. F. O'Neil, Jr. and M. Drillings, Eds., Erlbaum, Hillsdale, NJ. pp. 141–170.
- Tannenbaum, S. I., and Yukl, G. (1992), "Training and Development in Work Organizations," *Annual Review of Psychology*, Vol. 43, pp. 399–441.
- Teather, D. C. B., and Marchant, H. (1974), "Learning from Film with Particular Reference to the Effects of Cueing, Questioning, and Knowledge of Results," *Programmed Learning and Educational Technology*, Vol. 11, pp. 317–327.
- Thiagarajan, S. (1980), *Experiential Learning Packages*, Educational Technology Publications, Englewood Cliffs, NJ.
- Thorndike, R. L. (1986), "The Role of General Ability in Prediction," *Journal of Vocational Behavior*, Vol. 29, pp. 332–339.
- Thorndike, E. L., and Woodworth, R. S. (1901), "The Influence of Improvement in One Mental Function upon the Efficiency of Other Functions," *Psychological Review*, Vol. 8, pp. 247–261.
- Thorpe, J. A. (1987), "The New Technology of Large Scale Simulation Networking: Implications for Mastering the Art of Nonfighting," in *Proceedings of the 15th Annual Interservice/Industry Training Systems and Education Conference* (Washington, DC), pp. 87–98.
- Travers, R. M. (1967), "Research and Theory Related to Audiovisual Information Transmission," U.S. Office of Education contract No. OES-16-006, Western Michigan University, Kalamazoo (ERIC Document Reproduction Service No. ED 081 245).
- Tullar, W. L. (1989), "Relational Control in the Employment Interview," *Journal of Applied Psychology*, Vol. 74, pp. 971–977.
- Tulving, E., and Thomson, D. M. (1973), "Encoding Specificity and Retrieval Processes in Episodic Memory," *Psychological Review*, Vol. 3, pp. 112–129.
- Underwood, B. J. (1966), *Experimental Psychology*, 2nd Ed., Appleton-Century-Crofts, New York.

- Underwood, B. J., and Keppel, G. (1963), "Retention as a Function of Degree of Learning and Letter-Sequencing Interference," *Psychological Monographs*, Vol. 77, Whole No. 567.
- Vaill, P. B. (1971), "OD: A Grammatical Footnote," *Journal of Applied Behavioral Science*, Vol. 2, No. 2, p. 264.
- Vaill, P. B. (1989), *Managing as a Performing Art*, Jossey-Bass, San Francisco.
- Valverde, H. H. (1973), "A Review of Flight Simulator Transfer-of-Training Studies," *Human Factors*, Vol. 15, pp. 510–523.
- Watkins, M. J., and Tulving, E. (1975), "Episodic Memory: When Recognition Fails," *Journal of Experimental Psychology, General*, Vol. 1, pp. 5–29.
- Welford, A. T. (1968), *Fundamentals of Skill*, Methuen & Co, London.
- Welsh, J. R., Watson, T. W., and Ree M. J. (1990), *Armed Services Vocational Aptitude Battery (ASVAB): Predicting Military Criteria from General and Specific Abilities*, Air Force Human Resources Laboratory, Brooks Air Force Base, TX.
- Weaver, J. L., Morgan, B. B., Jr., Hall, J., and Compton, D. (1993), "Team Decision Making in the Command Information Center: Development of a Low-Fidelity Team Decision Making Task for Assessing the Effects of Teamwork Stressors," Technical Report submitted to the Naval Training Systems Center, Orlando, FL.
- Weaver, J. L., Bowers, C. A., Salas, E., and Cannon-Bowers, J. A. (1995), "Networked Simulations: New Paradigms for Team Performance Research," *Behavior Research Methods, Instruments, and Computers*, Vol. 27, No. 1, pp. 12–24.
- Wightman, D. C., and Sistrunk, F. (1987), "Part-Task Training Strategies in Simulated Carrier Landing Final-Approach Training," *Human Factors*, Vol. 29, pp. 245–254.
- Wixted, J. T., and Ebbesen, E. B. (1991), "On the Form of Forgetting," *Psychological Science*, Vol. 2, pp. 409–415.
- Zemke, R., and Kramlinger, T. (1982), *Figuring Things Out: A Trainer's Guide to Needs and Task Analysis*, Addison-Wesley, Reading, MA.