RAND

Designing Assessments of Microworld Training for Combat Service Support Staff

Susan G. Straus, Valerie L. Williams, John R. Bondanella, David Oaks, Emile Ettedgui

Prepared for the United States Army

Arroyo Center

Approved for public release; distribution unlimited

The research described in this report was sponsored by the United States Army under Contract No. DASW01-01-C-0003.

Library of Congress Cataloging-in-Publication Data

Designing assessments of microworld training for combat service support staff / Susan Straus.
p. cm.
"MR-1759."
Includes bibliographical references.
ISBN 0-8330-3448-0 (pbk. : alk. paper)
1. United States. Army—Supplies and stores. 2. Logistics—Study and teaching—United States—Evaluation. I. Straus, Susan.

UC263.D5 2003
355.4'11'071073—dc22

2003014993

Published 2003 by RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
201 North Craig Street, Suite 202, Pittsburgh, PA 15213-1516
RAND URL: http://www.rand.org/
To order RAND documents or to obtain additional information, contact Distribution Services: Telephone: (310) 451-7002; Fax: (310) 451-6915; Email: order@rand.org

Preface

This report describes new assessment alternatives for Army Combat Service Support (CSS) staff training. It was developed as part of a larger project to design training alternatives for the Army's Theater Support Command, which was newly formed at the time the project was initiated. The work was conducted in the Manpower and Training Program of RAND's Arroyo Center, a federally funded research and development center sponsored by the United States Army.

For more information on the Arroyo Center at RAND, contact the Director of Operations (telephone 310-393-0411, extension 6419; FAX 310-451-6952; e-mail Marcy_Agmon@rand.org), or visit the Arroyo Center's Web site at http://www.rand.org/ard/.

Contents

Figures

Tables

Summary

We propose new methods for evaluating learning outcomes from Combat Service Support (CSS) microworld training. RAND developed and piloted microworld training for distribution management skills as a part of a larger project that entailed making changes to the current structure, content, and methods of CSS training. Microworld models are small-scale simulations of organizations and operations. They are useful for training distribution management processes because they give the learner an opportunity to postulate changes and rapidly simulate the modified or new processes. Immediate feedback helps trainees understand how their actions affect the operation and how system components are interrelated. Consequently, trainees can learn proactive management of assets with realistic planning horizons and develop an understanding of the dynamic complexity affecting the organization.

We conducted pilot studies using a microworld model for U.S. Army Reserve (USAR) soldiers in Distribution Management Centers. The degree to which trainees learned training content was measured with a knowledge test in a pretest-posttest design, including a control group that received only the training and posttest. Results showed a statistically significant increase in trainees' knowledge of distribution management processes, although the gains appeared to be modest. The observed small increment in performance could have been due to the content and structure of the test, which may not have fully captured learners' knowledge or the factors that account for learning. Consequently, we propose a comprehensive set of evaluation measures to better assess learning and pinpoint areas for improvement in training content and processes. In this report, we present the following recommendations for the USAR and Combined Arms Support Command (CASCOM): (1) use a multidimensional, objective approach to measuring learning outcomes that includes cognitive, skill-based, and affective measures at multiple points in time; (2) provide training and collect learning measures via personal computer to facilitate training on demand, efficient data collection, and rapid feedback; and (3) provide incentives for the CSS centers and schools with responsibility for a function to help the field units develop training and assess training effectiveness. These recommendations should have wide application to a variety of process-oriented training programs in the field and in the schoolhouse.

Acknowledgments

The authors appreciate the contributions to this research by our sponsors and colleagues. Our co-sponsors—LTG Thomas Plewes, Chief, U.S. Army Reserve, and Mr. Thomas Edwards, Deputy to the Commanding General, U.S. Army Combined Arms Support Command—have provided support and insight throughout all phases of this project.

We wish to thank Dr. Matthew Lewis and Derek Davison, who contributed to the assessments described in this document, Nikki Shacklett, for her expert editorial support, and Vi-Nhuan Le, for her insightful review of the manuscript. Dr. John Winkler, while serving as Associate Director of the Arroyo Center's Manpower and Training Program, was especially involved in helping us focus on the most critical aspects of the Army Reserve training environment. Most of all, we are indebted to the Army Reserve soldiers—too many to name—who made significant suggestions for improving the nature of training that the Army will provide in the future.

Abbreviations

C2	Command and control
CASCOM	Combined Arms Support Command
CSS	Combat Service Support
DMC	Distribution Management Center
KS	Knowledge and skills
OPORDER	Operations order
PAM	Pamphlet
PC	Personal computer
TRADOC	Training and Doctrine Command
TSC	Theater Support Command
USAR	U.S. Army Reserve

1. Background

As the Army begins to reshape itself into the "Transformation Army," its ability to deploy quickly and conduct missions away from its garrison location places increasing importance on effective combat service support (CSS) command and control (C2). Previous Arroyo Center research argued that the structure, content, and methods used to train high-level CSS staff were insufficient for the future (Bondanella et al., 1999).

Specifically, the research noted that with regard to structure,

- training exercises are not integrated during a training year or across training years;
- training content is focused too narrowly on sustainment aspects of an operation only after the theater has matured;
- methods used for training are not intensive and realistic enough; and
- it is costly to maintain and run exercises.

Accordingly, the Arroyo Center developed and piloted a process-oriented approach to Army CSS training that entailed modifying the current structure, content, and methods of training. This process-oriented approach focused on the following goals:

- design an integrative structure in which the learning goals from one exercise lead into the goals for the next one;
- create training content that emphasizes proactive management of assets with realistic planning horizons; and
- develop a training methodology that uses microworld models to train distribution management processes.

One of the lessons we learned from this research is that determining the effectiveness of training is a complex process. As we explicate below, it was difficult to gauge how well trainees acquired the desired knowledge. Subsequently, we have identified theoretically- and empirically-derived methods for assessing whether training produces learning. The purpose of this report is to describe these methods and show how they can be applied to CSS C2 microworld training. In appendixes we provide a comprehensive, detailed set of measures that can be adapted to a variety of courses in CSS and other subjects.

We begin by summarizing the results of our previous research. Next, we interpret these findings and propose alternative methods for assessing learning outcomes. We conclude by addressing some of the issues related to developing and administering training assessments.

Microworld Models for Army Training

Microworld models are small-scale simulations of organizations and operations. Microworlds have been used successfully by the Arroyo Center as well as other organizations for training in a variety of contexts, including management, medicine, law enforcement, emergency management services, and aviation (Salas and Cannon-Bowers, 2001). Microworlds are useful for training different skills because they give the learner an opportunity to "learn by doing": the trainee can postulate changes in a system and then rapidly simulate the modified or new processes. Immediate feedback helps trainees understand how their actions affect the operation and how system components are interrelated. Research shows that training simulations facilitate skills transfer (Brannick, Prince, and Salas, unpublished manuscript, cited in Salas and Cannon-Bowers, 2001; Gopher, Weil, and Bareket, 1994). Microworlds that can be run on a network or personal computer are particularly appropriate for the Army's environment, which is characterized by rapid change. That is, the training content can be modified quickly, adapting to changes in tasks and technology. The training also can be conducted "any time, any place," providing training on demand and distributed learning.

Despite the suggested benefits of using microworld models for training, there are few reliable studies to assess their effects. In this report, we summarize the results of an earlier study by the Arroyo Center (Levy et al., 2001) that examined the effectiveness of using microworlds to train USAR soldiers in the design and management of logistics systems. The curriculum (described in Ettedgui, Oaks, and Bondanella, 1999) includes several focused (four-hour) seminars throughout the year, culminating in an annual two-week training exercise. Levy et al. focused on the first of these seminars.

The training seminar was designed around a microworld model that represented a simplified version of the distribution management process. The session's primary goals were to provide effective training in an environment with high personnel turnover, highlight the dynamic complexity of the organization, evaluate the impact of long time horizons on the design and management of theater distribution, and illuminate the effects of resource allocations and policy alternatives. Participants were trained in three broad categories of skills considered to be central to the performance of a Theater Support Command: basic system knowledge, measurement and trend identification, and course-of-action analysis. The training content was derived from the Materiel Distribution Management Skills Matrix described in Ettedgui et al. (1999) (see Figure 1). The study included both an experimental group, which received a pretest, training, and posttest, and a control group, for which the pretest was eliminated.[1] The test measured trainees' knowledge and their beliefs about the training effectiveness. Several of the knowledge questions were open-ended: for instance, one question presented a scenario and asked trainees to list the factors to consider in deciding whether to switch from a hub-and-spoke to a direct-delivery system.

Basic System Knowledge		**Measuring/Diagnosing, Managing Complex Systems**		
Basic/Enabling Knowledge/Skills	***System Understanding***	***Building/Expanding Operations***	***Executing Normal Operations of Systems***	***Handling Exceptions/ Crises***
Identify main supply routes	Develop/implement transportation networks	Design distribution network	Monitor status of distribution network and manage its flow	Take appropriate actions to work around chokepoints
Determine stockage levels	Understand alternative modes of distribution	Anticipate logistical requirements	Manage upstream supply/materiel flows	Prioritize requirements, reallocate resources to meet objectives
Understand modal resources	Understand host nation support, integration	Implement distribution network	Manage and track asset status	Integrate available assets to meet distribution needs
Filter/utilize logistics information	Understand automation information systems	Develop system that continually reviews relevant information	Keep system in equilibrium	Use alternative information-gathering techniques
Understand evolving logistics techniques, doctrine, and procedures	Understand just-in-time resource management	Incorporate state-of-the-art processes	Monitor relationships with sources of supply	Establish procedures to minimize crisis impact

Figure 1—Materiel Distribution Management Skills Matrix

Results showed that participants reacted favorably to the training and felt that their understanding of key concepts increased. Scores on the knowledge test also showed statistically significant improvement from pretest to posttest for questions that measured the ability to identify trends in data and understand the effects of those trends. However, the degree of knowledge was modest. For instance, the average posttest score for identifying trends was 1.84 on a 5-point scale. Questions that tested the ability of participants to analyze a course of

[1] The control group was included to examine if the pretest influenced performance on the posttest. There were no testing effects. The remaining discussion of results is based on the experimental group.

action showed an average posttest score of 1.17 on a 4-point scale. In addition, the absolute scores on the posttest were not consistent with participants' self-assessments of knowledge. Participants rated their understanding of key concepts at a level of 6–7 (on a 10-point scale) after training, a moderate-to-high assessment not warranted by the posttest scores.

One explanation for these results is that we should not expect learners to acquire in-depth knowledge of distribution management processes after completing the first session in a year-long program. However, an alternative explanation is that the test did not fully capture dimensions of student learning. The test focused narrowly on the CSS concepts addressed in the course. Although this is typical of training assessments, there are many other indicators of learning as well as factors that account for learning, as we delineate below. For example, students may not have had sufficient confidence or motivation to perform well on the knowledge test, or they may not have had the cognitive ability to master the course concepts. Thus, the test content and structure may have limited our ability to determine whether the trainees had successfully learned the skills we were teaching, and if not, how to revise the training sessions to increase their effectiveness.

The format of the test presents an additional issue in light of the Army's environment. The test consisted largely of open-ended questions, which require substantial training of test evaluators to ensure reliability and validity in scoring. We believe that a more efficient approach is needed to evaluate training effectiveness on an ongoing basis. It is particularly important to provide a method of testing that is more objective and less dependent on individual judgments in the Army's characteristic environment of frequent personnel turnover. A more objective method will enhance training efficiency, which is a key quality assurance goal of Army training (TRADOC PAM 350-70-4).

We address the role of testing in training in the remainder of this report by proposing alternative methods for evaluating learning from training within the context of using microworlds. In the following section, we describe our goals for testing and delineate a number of measures to assess trainees' knowledge. Our approach addresses the multifaceted nature of learning and the need for efficient methods of assessment.

2. Assessing Learning From Training

Objectives and Scope

Evaluation is a fundamental process in the Systems Approach to Training and contributes to the success of Army education and training (PAM 350-70-4). Our focus is on student learning, which is a basic component of internal evaluation of Army instructional systems. The primary purpose of our proposed evaluation scheme is to assess the degree to which soldiers learn logistical systems design and management from microworld training. In addition to showing whether soldiers have learned the material, incorrect responses should be an indicator of where the training fell short. Thus, the evaluation results should help improve future training sessions.

Many training evaluations consist of a limited range of measures. In field units, training evaluation is informal, consisting largely of after-action reviews. In classroom training, learning typically is assessed with unidimensional variables or methods. For instance, training evaluation often relies exclusively on multiple-choice knowledge tests. Although such tests are important, there are many other indicators of learning. A comprehensive approach to evaluating training requires a multidimensional approach. A broad range of measures can explain more of the variance in training outcomes and help training developers and instructors determine both whether trainees learn from training and the factors that explain why trainees do or do not learn.

A second goal is to use methods of assessment that are efficient. This objective is important in light of the rapidly changing environment in the Army, which includes changes in tasks, technology, and personnel.

Theoretical Framework and Proposed Measures

Kraiger, Ford, and Salas (1993) proposed a framework for evaluating training that includes three categories of learning: cognitive outcomes, skill-based outcomes, and affective outcomes. Examples of cognitive outcomes include declarative knowledge (i.e., information about "what"), procedural knowledge (information about "how"), and knowledge organization (cognitive maps or mental models) (Gagne, 1984). Examples of skill-based outcomes include skill acquisition, compilation (evidenced by smooth, fast performance with few errors), and automaticity (when the learner no longer has to give conscious effort to each step in a sequence of behaviors). Affective outcomes include attitudes to perform well, motivation to learn the task, levels of performance goals, and self-efficacy (confidence in one's ability to perform a task). Research on microworlds has shown that all three categories of learning outcomes predict training effectiveness (Ford, Smith, Weissbein, Gully, and Salas, 1998; Kozlowski et al., 2001). We use the Kraiger et al. model as a foundation to propose how one might evaluate learning in microworld training of logistical systems design and management.

Table 1 presents proposed learning outcomes and measurement methods. We do not include all of the measures delineated by Kraiger et al., as some variables would not be practical to measure in the context of CSS training (e.g., "automaticity"). Following Table 1, we describe the rationale for selecting these outcomes and measures and provide examples of items tailored to the CSS C2 course. We also discuss when to administer the measures, as the information gained by different measures may depend on the stage of training.

Table 1

Classification of Learning Outcome Measures

Category	Learning Outcome	Measure
Cognitive	Verbal knowledge —Declarative knowledge —Procedural knowledge	Multiple choice test items (see Appendix A)
	Knowledge organization	Paired comparisons of training concepts
	Cognitive strategies	Self-assessments of knowledge (Appendix B)
Skill-based	Skill-acquisition and compilation	Perform 5–10 trials of computer-based microworld
	Compilation	Perform additional trials of the microworld set in a different context
Affective	Motivation to learn	Self-report (see Appendix C)
	Mastery and performance orientations	Self-report (see Appendix D)
	Self-efficacy	Self-report (see Appendix E)
Control variables	e.g., Age, rank, education, cognitive ability	Self-report or from personnel records

Cognitive Learning Outcomes

One component of cognitive learning is verbal knowledge, or information that one can articulate. Two fundamental types of verbal knowledge include declarative knowledge and procedural knowledge. Declarative knowledge refers to information such as facts, principles, theories, and opinions. Examples of declarative knowledge in the context of CSS training include the ability to define terms such as a "node" and identify patterns such as "starving the customer" or "growing to the maximum." Procedural knowledge refers to knowledge of **how** to perform tasks or cognitive activities. Procedural knowledge is a higher-order level of knowledge; it reflects how to bring declarative knowledge to bear in performing a task and the ability to generalize that process to different conditions. An example of procedural knowledge is knowing the steps to take in distribution design and management (e.g., the first step is to perform a CSS mission analysis).

Declarative and procedural knowledge can be assessed via the amount of knowledge, the accuracy of recall, or the speed of recall. We propose using multiple-choice items to assess these types of knowledge. Such measures of learning have been used with increasing frequency in research on training (e.g., Ford et al., 1998; Kozlowski et al., 2001). Appendix A provides examples of items for a multiple-choice test that measures verbal knowledge for CSS microworld training. For instance, an item that tests declarative knowledge is: Bottlenecks indicate a breakdown in which of the following components of distribution management? (a) control; (b) reliability; (c) visibility; and (d) capacity. An example of an item that measures procedural knowledge is: Which of the following is a correct progression of steps in distribution design and management? (e.g., (a) Estimate resources, analyze combat plan, monitor execution, publish OPORDER). Items would be administered on appropriate content following each of the three focused training seminars in the curriculum.

Regarding the timing of measurement, it is particularly important to measure declarative knowledge early on. First, the acquisition of declarative knowledge occurs in initial stages of training. Second, as noted earlier, declarative knowledge is necessary, but not sufficient, for higher-order skill development (e.g., Kozlowski et al., 2001). Evidence that trainees have not acquired declarative knowledge would suggest that they might have difficulty acquiring procedural knowledge. In addition, there is likely to be greater variability in declarative knowledge early versus late in training. Therefore, earlier scores can be used for predicting other learning outcomes. However, verbal knowledge should be measured at the conclusion of training as well. It is also useful to measure these outcomes some time after training has occurred to assess knowledge retention (Wisher, Curnow, and Seidel, 2001).

Knowledge organization is another cognitive learning outcome. The term *mental model* refers to how people organize knowledge. For instance, studies show that novices and experts have different mental models for the same task and that the complexity of these models is associated with performance. When compared with those of novices, experts' mental models show closer links between problems and solutions. Consequently, experts can access solutions to problems more quickly (Glaser and Chi, 1989).

Mental models can be measured with rating tasks, such as paired comparisons, in which trainees evaluate the degree of similarity among training concepts (e.g., Mathieu, Heffner, Goodwin, Salas, and Cannon-Bowers, 2000). Cluster analysis or network analysis is used to identify respondents' mental models and demonstrate convergence with experts' ratings of the same concepts. The more the trainees have learned, the more their responses will map on to those of the experts. Alternatively, trainees' ratings can be analyzed for coherence, or consistency of knowledge structures (Kozlowski et al., 2001). For instance, if A is related to B, and B is not related to C, then A should not be related to C (Kozlowski et al., 2001). Ideally, one would administer the comparison task to trainees at several points in training; increasing similarity to experts' responses or increased coherence would indicate training effectiveness, and a lack of improvement would suggest that trainees are not acquiring the "big picture."

To measure mental models in CSS microworld training, we would administer a paired-comparison task to subject matter experts to establish a standard. Trainees would perform the same rating task following training. Examples of distribution management concepts that participants might evaluate include node, hub, spoke, port, bottleneck, staging, and integration. If time permitted, we would administer the paired-comparison task after each focused seminar and after the final two-week training exercise to examine changes in responses over time.

Cognitive strategies reflect the speed with which knowledge can be applied. As trainees develop higher-order skills, they gain more sophisticated strategies for performing the task. Typically, methods of measuring this outcome are too labor intensive to meet our goal for efficiency in assessment (e.g., individual protocol analysis in which each trainee is asked to verbalize the steps required to perform a task and the reasons for taking those actions). However, some researchers have found that trainees' self-assessments of learning are reasonably accurate indicators of their knowledge (e.g., Fisk and Gallini, 1989; Slife and Weaver, 1992). As noted in Wisher and Curnow (1998), research also has reported significant, positive correlations between self-assessments and job performance or related measures. Self-assessments can be collected in an efficient manner. Examples from previous research include: "How many items do you expect to answer correctly on a 50-item test of this course material?" (e.g., Fisk and Gallini, 1989) and "How many practice sessions will you need to score: 50% or better?; 75% or better?; 95% or better?" (e.g., Schendel and Hagman, 1982). For CSS training, we would use an adaptation of the self-assessment administered in the original Arroyo Center CSS microworld training study reported in Levy et al. (2001) (see Appendix B). For example, trainees are asked to rate their understanding of concepts, such as theater support command organizational structure, and processes, such as how to evaluate alternative distribution policies. We would administer the self-assessment before the other tests, such as the multiple-choice exam and skill-based assessments (described in the next section).

Skill-Based Learning Outcomes

Skill-based outcomes include skill acquisition, compilation, and automaticity. Skill acquisition is the development of technical or motor skills. It reflects the ability to enact procedural knowledge (i.e., doing versus knowing). Compilation is characterized by faster and more fluid performance, without the need for mentally rehearsing procedures or routines. When performance reaches an automatic state, conscious monitoring is no longer required, and the individual can perform other tasks simultaneously. As noted earlier, however, automaticity may be difficult to assess in some types of training; furthermore, soldiers may be expected to develop automaticity on the job rather than in classroom training.

Multiple-choice tests can be used to assess whether trainees know the steps to take to perform a task (e.g., see Appendix A), but behavioral observation is required to determine whether trainees can perform the task correctly. Typically, one assesses skill-based outcomes by observing trainees in role-plays during training or performing their jobs after the training. Whereas it may not be practical to conduct systematic on-the-job observations in the field, microworld simulations provide an ideal context for evaluating skill acquisition and compilation during training. For instance, after having the opportunity to practice skills using the simulation (e.g., after the third seminar in focused staff training), trainees might complete several (e.g., 5–10) trials of the simulation for testing purposes. Performance on these trials would reflect the trainee's ability to formulate an effective course of action in distribution management. A variety of performance measures could be used, such as: (1) the incidence of errors such as underutilization of materiel, bottlenecks, and overstocked nodes; (2) the magnitude of these errors; (3) time to solution; and (4) the extent to which error rates and speed improve across trials. The type and incidence of errors would indicate which aspects of the task trainees have or have not learned.

This test can assess both initial skill acquisition and compilation. Typical measures of compilation include behavioral observations and hands-on performance tests using trained observers, and situational interviews in which students describe the steps to take to complete a task. These methods are labor intensive, requiring one-on-one testing. Whereas there are some tasks for which individual hands-on assessment is essential, we recommend finding alternative approaches where feasible. Accordingly, one indication of compilation is the ability to generalize behaviors to new situations. We can test this outcome with minimal resources by having trainees perform additional trials of the simulation in a different context (e.g., managing the logistics of personnel versus ammunition). Scores would indicate the extent to which trainees have grasped both the situation-specific and general concepts that apply to distribution management.

Affective Learning Outcomes

There are a variety of affective measures that are associated with learning during training and application of skills following training. Affective outcomes include trainees' attitudes and motivation. These measures are distinct from reactions to training (e.g., perceptions of the quality of instruction), which provide feedback on aspects of training delivery but do not assess learning. Some affective outcomes (e.g., attitudes and self-efficacy) are direct measures of learning in that one expects to see changes in these constructs from initial to later stages of training. Other affective outcomes (e.g., motivation to learn) are important because they help predict the extent to which trainees will learn substantive course material (e.g., low scores on a knowledge test could be due to a lack of motivation to learn, which in turn might be caused by an individual's perception that the training course will not improve his or her job performance). In fact, Gagne (1984) argues that failure to assess affective measures provides an incomplete picture of learning and the learning process.

Attitudes indicate the degree to which trainees identify with training goals. Measures showing a shift in attitudes from pre- to post-training (e.g., from apathy to commitment) can be an indicator of learning. However, these measures are typically used when training addresses value-laden topics, such as tolerance for diversity. As CSS training is concerned largely with learning factual concepts and procedures, we do not propose measures of attitude change in this course.

Motivational outcomes reflect the reasons that trainees want to participate in training and their beliefs about their own effectiveness at the task. Colquitt, LePine, and Noe (2000) report that motivation to learn is positively related to other training outcomes, including declarative knowledge, skill acquisition, post-training self-efficacy, and reactions to training. Appendix C presents items measuring motivation to learn adapted from Noe and Wilk (1993) and Noe and Schmitt (1986). Examples include: "I am willing to exert considerable effort in this course to improve my skills," and "Participating in training is of little use to me because I have all the knowledge and skills I need to successfully perform my job." These items would be administered prior to the first training session.

Another aspect of motivation that has gained increasing research attention is trainees' orientations toward performance versus mastery. Someone who holds a performance orientation is concerned about doing well and being perceived favorably by others. Someone who holds a mastery orientation shows more concern for increasing his or her competence on the task (Dweck and Leggett, 1988). These orientations are trait-based but can also be influenced by external cues or instructions, and both trait-based and externally mediated goals influence training outcomes (Kozlowski et al., 2001). Several studies show that trainees who have a mastery orientation show superior learning processes and outcomes, such as engaging in more effective learning strategies and exhibiting more coherent mental models, better learning outcomes, a more positive affect toward the task, and high self-efficacy for the task (Ford et al., 1998; Kozlowski et al., 2001; Stevens and Gist, 1997). A scale developed by Button, Mathieu, and Zajac (1996) is the most widely used measure of training orientation in research (see Appendix D). For instance, an item measuring a mastery orientation is: "I prefer to work on tasks that force me to learn new things." An example of a performance orientation item is: "I prefer to do things that I can do well rather than things that I do poorly." Motivational orientation would be measured prior to training.

Self-efficacy is another motivational outcome that has implications for learning. Self-efficacy refers to the belief that one is capable of performing a particular task. Numerous studies demonstrate the association of self-efficacy with other training outcomes. For instance, both Ford et al. (1998) and Kozlowski et al. (2001) found that higher self-efficacy was associated with superior transfer-of-training. Training can create and strengthen self-efficacy (Bandura, 1977). Consequently, an increase in self-efficacy from the early to the late stages of training may be an indication of learning. Self-efficacy measured late in training also can predict transfer of training and long-term job performance on the task in question. A common measure of self-efficacy presents task performance outcomes of increasing difficulty and asks respondents to indicate how confident they are that they could perform each outcome on a scale from 1 to 10. Self-efficacy is measured by the average confidence rating across all items (Stevens and Gist, 1997).

Examples of self-efficacy questions for CSS training were derived from the Materiel Distribution Management Skills matrix (Figure 1) and are presented in Appendix E. An example is: "I could predict an oversupply of materiel at a node: 2 days before it occurred; 4 days before it occurred; 6 days before it occurred; 8 days before it occurred; 10 days before it occurred." Respondents are asked to rate their confidence of achieving each performance level on a 10-point scale. We would measure self-efficacy prior to the first seminar, after the last seminar, and after the two-week focused exercise.

Control Variables

Of course, factors such as level of education and other demographic variables can affect test performance, as reported in Levy et al. (2001). Consequently, we suggest collecting demographic data including the trainee's rank, unit and unit tenure, job position and job tenure, specialty code, military and civilian educational background, age, experience, prior relevant training, and general cognitive ability (as represented by the Armed Forces Qualification Test score). These data can be used to examine the extent to which demographic characteristics and cognitive ability account for performance on measures of learning. These data can be obtained from trainees at the beginning of training or through archival sources (personnel records).

3. Reactions

Reactions to training consist of affective reactions and utility reactions. Affective reactions refer to how much participants liked training (e.g., "I enjoyed participating in this training session."). Affective reactions are the most common training outcome measure, but they are not associated with learning or job performance (Alliger, Tannenbaum, Bennett, Traver, and Shotland, 1997). Nonetheless, affective reactions may be important to the extent that they predict whether trainees will enroll in other courses or encourage peers to participate. Consequently, we suggest measuring affective reactions but including only a limited number of items (see Appendix F). Utility reactions refer to how useful participants found the training. This construct is encompassed by some of the motivation-to-learn items presented in Appendix C, such as "Participating in this course will help me to advance my career."

4. Measurement Implications

A multidimensional assessment approach can provide a variety of information about how to improve training effectiveness. For example, some results might indicate that instructors should place additional emphasis on particular topics. Alternatively, results might show that substandard test performance is due to student characteristics rather than training content and structure. In this section, we describe how these measures might be used to guide efforts to improve training.

Cognitive and Skill-Based Outcomes

Tests of knowledge and skills (KS), including multiple-choice tests of declarative and procedural knowledge, self-assessments of learning, and trials on the microworld simulation, serve a number of purposes. Scores on the KS tests and microworld trials can be used to determine passing rates or go/no go decisions. From a diagnostic perspective, responses to questions or tests can indicate which aspects of training students have or have not mastered. For instance, incorrect responses on the knowledge test, poor performance on criteria in the microworld trials, or low ratings on the self-assessment of learning may point to specific topics that need to be addressed in more depth during training. Satisfactory performance on the microworld trials in the focal training area (e.g., water distribution), but not in other domains (e.g., personnel), may suggest that students grasp the situation-specific concepts but not the general ones. In this case, students may need more practice in diverse contexts to improve their ability to generalize the course concepts. Mental model scores that do not bear increasing resemblance to those of subject matter experts over the course of training would indicate that students are not learning the appropriate connections among course concepts.

Affective Outcomes

Affective outcomes also can be used for diagnostic purposes. Scores on the motivation-to-learn scales can be used in a number of ways. First, if scores on the utility of training items (14a–h in Appendix C) are low at the onset of training and do not increase significantly over time, this suggests that students have low expectations about the value of training and that participating in the course does not change their opinions. The correlation between motivation scores and KS tests also indicates whether motivation for training predicts performance. For example, a positive correlation between motivation scores and KS test scores indicates that students who do not plan to exert a lot of effort in the course or do not see the value of the course perform worse. Consequently, one could select students for training based on their motivation or revise the program of instruction to highlight the benefits of the course.

Similarly, the association between motivational orientation scores and performance on KS tests would indicate whether having a performance or mastery orientation predicts training performance. For instance, one might expect students with a mastery orientation to perform significantly better than those with a performance orientation. If this prediction is supported, it suggests selecting students for training based on their motivational orientation score or creating a mastery orientation via cues from instructors.

Self-efficacy scores that are low at the beginning of training (which one would expect) and that fail to increase significantly would suggest that students do not believe they can perform the task effectively, despite training. If post-training KS test scores are also low, then the students' insecurity may be warranted, and improvements may be needed in training content or structure. But if KS test scores are satisfactory, then an intervention that addresses student confidence would be needed.

Variables such as demographic characteristics serve as control factors. For instance, controlling for age or rank when analyzing the association between motivation and performance can provide a more precise correlation between motivation and performance. Potentially, these variables also could be used to select soldiers for training. For instance, analyses that control for the effect of experience or unit tenure on KS performance might show that these factors are significant predictors of training performance. These results would suggest that students must meet a minimum threshold of experience or tenure to participate in training.

5. Mode of Administration

In addition to identifying appropriate measures of learning outcomes, we propose administering all measures by computer, thereby integrating the training itself with the evaluation of training. Trainees would be able to access training and assessment modules on a CD-ROM or over a network using existing unit resources. Computer (versus paper-and-pencil) administration will improve accuracy and efficiency in data collection and scoring and provides a mechanism to give rapid feedback to trainees, instructors, and training developers. This mode of administration does not require participants to be computer specialists; it only requires that they are adept at using PCs. Using web-based delivery also can reduce the need for external personnel support. Changing the method of training and test administration can facilitate training at a distance and learning on demand.

6. Training and Assessment Development

Our research has focused on training in field units; it is not institutional training. Whereas personnel in the field are responsible for delivering training, they do not have the background or resources to identify or develop the measures required to assess training effectiveness without substantial assistance from the institutional training base. Thus, we recommend that the CSS centers and schools with responsibility for a function should assist the field units in assessing training effectiveness. In addition, by using a web-based approach, CSS centers and schools can provide measures that unit trainers can access easily to conduct systematic assessments of unit training. The USAR and CASCOM can facilitate this process further by reallocating funding and personnel resources to unit training.

7. Conclusions

We have described a multidimensional approach to measure learning from CSS C2 microworld training. Recommendations for training assessments are based on recent theoretical developments and empirical findings in the training literature and are framed within the context of the training year plan outlined in Bondanella et al. (1999). We also offer recommendations about the mode of training and test administration and emphasize the importance of involving the institutional training base in training and assessment development for units. Taken together, these recommendations should help provide more effective training for field units as the Transformation Army evolves.

The recommendations we offer in this report have broad implications for training and assessment design. For instance, microworlds and other simulations are appropriate to a variety of training topics that deal with system dynamics of combat service support of military operations in field environments. Such topics include deployment of organizations from disparate locations to multiple bases in a contingency area, individual personnel replacement, medical services, and maintenance. The multidimensional approach to assessing training effectiveness also applies to nearly any training topic and can be used in field units (with appropriate support) or residence training. Of course, the content of specific measures would need to be adapted to factors such as the topic and structure of training. For example, the content of a knowledge test and the items used in paired comparisons would be different for distribution management versus medical services. Likewise, the use of personal computers for training delivery and assessment has applications to a wide variety of training topics. Training via personal computer or the Internet also is the foundation of distance learning. As noted by Winkler, Leonard, and Shanley (2001), distance learning can offer numerous benefits to the Army, including alleviating shortages of qualified enlisted personnel in critical jobs and fostering the stability and development of soldiers and leaders. We encourage the Army to implement approaches to training and assessment consistent with our recommendations to enhance learning and performance in an environment characterized by rapid change.

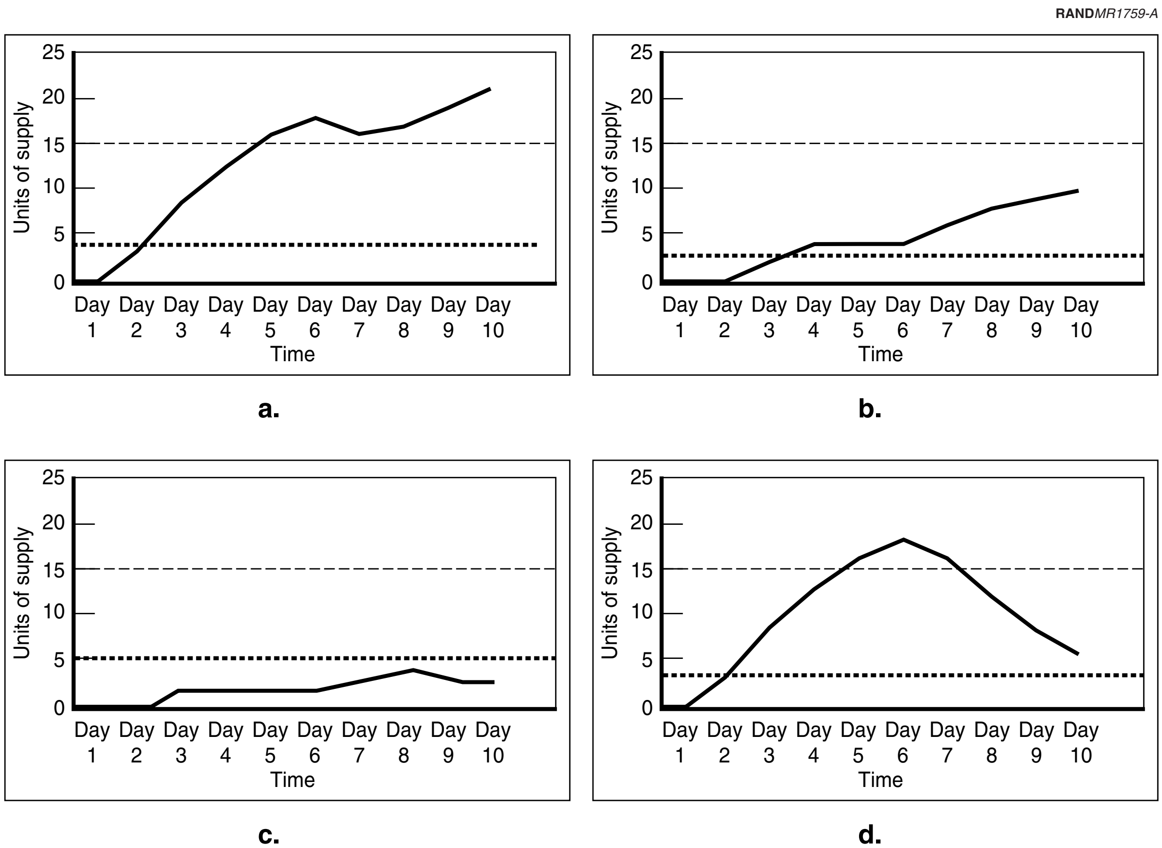
Appendix

A. Sample Multiple Choice Knowledge Test Items

INSTRUCTIONS: The following questions test concepts from distribution management training. Circle the letter corresponding to the best answer for each question. Choose only one response for each question.

1. The most essential component of distribution management is:
   a. control
   b. reliability
   c. visibility
   d. capacity
2. Bottlenecks indicate a breakdown in which of the following components of distribution management?
   a. control
   b. reliability
   c. visibility
   d. capacity
3. Which of the following principles of distribution management is INCORRECT?
   a. minimize throughput, maximize stockpiling
   b. centralize management at each echelon
   c. optimize the Theater Distribution System
   d. substitute velocity for mass and reduce logistics response time
4. Which of the following is an example of the **physical network** component of infrastructure?
   a. roads
   b. host nation support
   c. distribution assets
   d. information technology

5. Consider the following four diagrams. Which pattern illustrates an overstocked node?



6. Which of the following is a correct progression of steps in distribution design and management? (note: these are a subset of the total number of steps needed)

   a. Estimate resources, analyze combat plan, monitor execution, publish OPORDER

   b. Publish OPORDER, monitor execution, perform CSS mission analysis, estimate resources

   c. Validate distribution plan, publish OPORDER, monitor execution, analyze combat plan

   d. Perform CSS mission analysis, estimate resources, publish OPORDER, monitor execution

Questions 7 and 8 are based on the following scenario. You may use the spreadsheets to work out the answers to the questions.

Scenario: You are a water supply officer in a newly opened, developing theater in an arid climate. It is your first priority to meet the potable water needs of the deploying force. The table below shows the anticipated counts of supported personnel in the theater at the end of each day for the first ten days.

You already have in theater a Reverse Osmosis Water Purification Unit (ROWPU) (fed from a large, brackish lake) and a distribution system that allow you to make and deliver up to 4,500 gallons of potable water a day.

7. Assuming that personnel in this theater need 9 gallons of potable water per day (per person), and assuming that you cannot store surplus water, on what day will you need an additional source of potable water?

Day	1	2	3	4	5	6	7	8	9	10
Personnel	200	300	400	450	500	600	700	800	900	1,000
Production quantity										
Consumption quantity										

a. 5
b. 6
c. 7
d. 8

8. Assuming that personnel in this theater need 15 gallons of potable water per day (per person), and assuming that you CAN store surplus water, on what day will you need an additional source of potable water?

Day	1	2	3	4	5	6	7	8	9	10
Personnel	200	300	400	450	500	600	700	800	900	1,000
Production quantity										
Consumption quantity										

a. 1
b. 2
c. 3
d. 4

B. Self-Assessment of Knowledge

(Adapted from Levy, Lewis, Bondanella, Baisden, and Ettedgui, 2001)

INSTRUCTIONS: Please rate your understanding of each of the following topics from this course. Circle one number on the scale, from 1 = No Understanding to 7 = Excellent Understanding.

Item	Response Options (circle one)						
	No Understanding			Moderate Understanding			Excellent Understanding
1. Theater Support Command (TSC) organizational structure	1	2	3	4	5	6	7
2. Distribution Management Center (DMC) organizational structure	1	2	3	4	5	6	7
3. DMC Mission Essential Task List	1	2	3	4	5	6	7
4. How the DMC will design theater distribution systems in general	1	2	3	4	5	6	7
5. How to measure the operation of a TSC	1	2	3	4	5	6	7
6. How to evaluate alternative distribution policies	1	2	3	4	5	6	7
7. The potential impacts of distribution policies over time	1	2	3	4	5	6	7
8. The potential impacts of distribution management policies developed in my unit on other organizations within the TSC	1	2	3	4	5	6	7
9. The potential impacts of my distribution management decisions on other units within the TSC	1	2	3	4	5	6	7
10. How to manage a general supply distribution network	1	2	3	4	5	6	7
11. How to manage a personnel replacement network	1	2	3	4	5	6	7
12. How to manage a POL (Petroleum, Oil and Lubricants) distribution network	1	2	3	4	5	6	7

C. Motivation to Learn

(Adapted from Noe and Schmitt, 1986, and Noe and Wilk, 1993)

INSTRUCTIONS: Please rate the extent to which you disagree or agree with each statement. Circle one number on the scale, from 1 = Strongly disagree to 7 = Strongly agree.

Item	Response Options (circle one) Strongly disagree						Strongly agree
1. I am willing to exert considerable effort in this course to improve my skills.	1	2	3	4	5	6	7
2. I believe that I can improve my skills by participating in this course.	1	2	3	4	5	6	7
3. I am motivated to learn the skills emphasized in this course.	1	2	3	4	5	6	7
4. I will try to learn as much as I can from this course.	1	2	3	4	5	6	7
5. I will get more from this training than most people will.	1	2	3	4	5	6	7
6. When I'm involved in training and I can't understand something, I get so frustrated I stop trying to learn.	1	2	3	4	5	6	7
7. I believe I can learn the material presented in the training program.	1	2	3	4	5	6	7
8. I would like to improve my skills.	1	2	3	4	5	6	7
9. My present job performance satisfies my personal expectations and goals.	1	2	3	4	5	6	7
10. Participating in training is of little use to me because I have all the knowledge and skills I need to successfully perform my job.	1	2	3	4	5	6	7
11. I am willing to invest effort to improve skills related to my current job.	1	2	3	4	5	6	7
12. I am willing to invest effort to improve skills just for the sake of learning.	1	2	3	4	5	6	7
13. I am willing to invest effort to improve skills in order to prepare myself for a promotion.	1	2	3	4	5	6	7

14. Participating in this course will help me to:	Strongly disagree						Strongly agree
a. Grow as a person.	1	2	3	4	5	6	7
b. Increase my self-confidence.	1	2	3	4	5	6	7
c. Obtain respect from my peers.	1	2	3	4	5	6	7
d. Obtain a salary increase.	1	2	3	4	5	6	7
e. Perform my job better.	1	2	3	4	5	6	7
f. Obtain the job I want.	1	2	3	4	5	6	7
g. Advance my career.	1	2	3	4	5	6	7
h. Acquire new knowledge.	1	2	3	4	5	6	7

D. Motivational Orientation

(Used in its entirety from Button, Mathieu, and Zajac, 1996)

INSTRUCTIONS: Please rate the extent to which you disagree or agree with each statement. Circle one number on the scale, from 1 = Strongly disagree to 7 = Strongly agree.

Item	Response Options (circle one) Strongly disagree						Strongly agree
1. I prefer to do things that I can do well rather than things that I do poorly.	1	2	3	4	5	6	7
2. I'm happiest at work when I perform tasks on which I know that I won't make any errors.	1	2	3	4	5	6	7
3. The things I enjoy the most are the things I do the best.	1	2	3	4	5	6	7
4. The opinions others have about how well I can do certain things are important to me.	1	2	3	4	5	6	7
5. I feel smart when I do something without making any mistakes.	1	2	3	4	5	6	7
6. I like to be fairly confident that I can successfully perform a task before I attempt it.	1	2	3	4	5	6	7
7. I like to work on tasks that I have done well on in the past.	1	2	3	4	5	6	7
8. I feel smart when I can do something better that most other people.	1	2	3	4	5	6	7
9. The opportunity to do challenging work is important to me.	1	2	3	4	5	6	7
10. When I fail to complete a difficult task, I plan to try harder the next time I work on it.	1	2	3	4	5	6	7
11. I prefer to work on tasks that force me to learn new things.	1	2	3	4	5	6	7
12. The opportunity to learn new things is important to me.	1	2	3	4	5	6	7
13. I do my best when I'm working on a fairly difficult task.	1	2	3	4	5	6	7

14. I try hard to improve on my past performance.	1	2	3	4	5	6	7
15. The opportunity to extend the range of my abilities is important to me.	1	2	3	4	5	6	7
16. When I have difficulty solving a problem, I enjoy trying different approaches to see which one will work.	1	2	3	4	5	6	7

E. Self-Efficacy Items

INSTRUCTIONS: For each question, rate how confident you are that you can achieve each level of performance *today*, without additional training. Circle a number from 1 = not at all confident to 10 = totally confident.

Below is an example in a different context, i.e., the task of running a mile within different time limits:

EXAMPLE:

RANDMR1759-E

Question	Confidence of achieving each level									
I could run a mile in:	Not at all Confident									Totally Confident
11 minutes	1	2	3	4	5	6	7	8	9	(10)
9 minutes	1	2	3	4	5	6	7	8	9	(10)
7 minutes	1	2	3	4	5	6	(7)	8	9	10
5 minutes	1	2	3	4	(5)	6	7	8	9	10
3 minutes	(1)	2	3	4	5	6	7	8	9	10

Question	Confidence of achieving each level									
1. I could explain how to identify main supply routes with:	Not at all Confident									Totally Confident
20% accuracy:	1	2	3	4	5	6	7	8	9	10
40% accuracy:	1	2	3	4	5	6	7	8	9	10
60% accuracy:	1	2	3	4	5	6	7	8	9	10
80% accuracy:	1	2	3	4	5	6	7	8	9	10
100% accuracy:	1	2	3	4	5	6	7	8	9	10
2. I could teach my counterpart from another TSC the concepts behind Just-In-Time resource management with:	Not at all Confident									Totally Confident
20% accuracy:	1	2	3	4	5	6	7	8	9	10
40% accuracy:	1	2	3	4	5	6	7	8	9	10
60% accuracy:	1	2	3	4	5	6	7	8	9	10
80% accuracy:	1	2	3	4	5	6	7	8	9	10
100% accuracy:	1	2	3	4	5	6	7	8	9	10

3. I could design a first draft of a distribution system for a small-size contingency if I had all the necessary information, staff, and OPORDER in:	Not at all Confident									Totally Confident
20 hours:	1	2	3	4	5	6	7	8	9	10
16 hours:	1	2	3	4	5	6	7	8	9	10
12 hours:	1	2	3	4	5	6	7	8	9	10
8 hours:	1	2	3	4	5	6	7	8	9	10
4 hours:	1	2	3	4	5	6	7	8	9	10
4. I could obtain a sufficient supply of potable water in 7 days for:	Not at all Confident									Totally Confident
100 personnel:	1	2	3	4	5	6	7	8	9	10
300 personnel:	1	2	3	4	5	6	7	8	9	10
500 personnel:	1	2	3	4	5	6	7	8	9	10
700 personnel:	1	2	3	4	5	6	7	8	9	10
900 personnel:	1	2	3	4	5	6	7	8	9	10
5. I could predict an oversupply of materiel at a node:	Not at all Confident									Totally Confident
2 days before it occurred:	1	2	3	4	5	6	7	8	9	10
4 days before it occurred:	1	2	3	4	5	6	7	8	9	10
6 days before it occurred:	1	2	3	4	5	6	7	8	9	10
8 days before it occurred:	1	2	3	4	5	6	7	8	9	10
10 days before it occurred:	1	2	3	4	5	6	7	8	9	10
6. I could foresee a node in the distribution network becoming a bottleneck system:	Not at all Confident									Totally Confident
2 days before it occurred:	1	2	3	4	5	6	7	8	9	10
4 days before it occurred:	1	2	3	4	5	6	7	8	9	10
6 days before it occurred:	1	2	3	4	5	6	7	8	9	10
8 days before it occurred:	1	2	3	4	5	6	7	8	9	10
10 days before it occurred:	1	2	3	4	5	6	7	8	9	10
7. I could foresee a "starving client" node in a distribution network:	Not at all Confident									Totally Confident
1 day before it occurred:	1	2	3	4	5	6	7	8	9	10
3 days before it occurred:	1	2	3	4	5	6	7	8	9	10
5 days before it occurred:	1	2	3	4	5	6	7	8	9	10
7 days before it occurred:	1	2	3	4	5	6	7	8	9	10
9 days before it occurred:	1	2	3	4	5	6	7	8	9	10

F. Affective Reactions

INSTRUCTIONS: Please rate the extent to which you disagree or agree with each statement. Circle one number on the scale, from 1 = Strongly disagree to 7 = Strongly agree.

Item	Response Options (circle one)						
	Strongly disagree						Strongly agree
1. I enjoyed participating in this course.	1	2	3	4	5	6	7
2. The course was interesting.	1	2	3	4	5	6	7
3. The instructors were knowledgeable about the course material.	1	2	3	4	5	6	7
4. The instructors were responsive to my questions.	1	2	3	4	5	6	7
5. I received sufficient feedback about my performance.	1	2	3	4	5	6	7
6. There was enough time in the course to learn the material.	1	2	3	4	5	6	7
7. There was enough time in the course to practice my skills.	1	2	3	4	5	6	7
8. The instructors motivated me to do my best.	1	2	3	4	5	6	7
9. I learned a lot from written course materials.	1	2	3	4	5	6	7
10. I learned a lot from the practical exercises.	1	2	3	4	5	6	7
11. I was satisfied with the equipment and supplies used in the course.	1	2	3	4	5	6	7
12. The course was taught at the right level—neither too easy nor too hard.	1	2	3	4	5	6	7

References

Alliger, G.M., Tannenbaum, S.I., Bennett, W., Traver, H., and Shotland, A. (1997). A meta-analysis of the relations among training criteria. *Personnel Psychology*, *50*, 341–358.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191–215.

Bondanella, J.R., Lewis, M.W., Steinberg, P., Shin-Kyu Park, G., Levy, D.G., Ettedgui, E., Oaks, D.M., Sollinger, J.M., Winkler, J.D., Halliday, J.M., and Way-Smith, S. (1999). *Microworld Simulations for Command and Control Training of Theater Logistics and Support Staffs: A Curriculum Strategy*, Santa Monica, CA: RAND, MR-929-A.

Button, S.B., Mathieu, J.E., and Zajac, D.M. (1996). The development and psychometric evaluation of measures of learning goal and performance goal orientation. *Organizational Behavior and Human Decision Processes*, *67*, 26–48.

Colquitt, J.A., LePine, J.A., and Noe, R.A. (2000). Toward an integrative theory of training motivation: A meta-analytic path analysis of 20 years of research. *Journal of Applied Psychology*, *85*, 678–707.

Dweck, C.S., and Leggett, E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–273.

Ettedgui, E., Oaks, D.M., and Bondanella, J.R. (1999). *The Use of Microworld Simulations to Train Theater Level CSS Staffs: Training Development Considerations*, Santa Monica, CA: RAND, DB-265-A.

Fisher, S.L., and Ford, J.K. (1998). Differential effects of learner effort and goal orientation on two learning outcomes. *Personnel Psychology*, *51*, 397–420.

Fisk, A.D., and Gallini, J.K. (1989). Training consistent components of tasks: Developing an instructional system based on automatic/controlled processing principles. *Human Factors*, *31*, 453–463.

Ford, J.K., Smith, E.M., Weissbein, D.A., Gully, S.M., and Salas, E. (1998). Relationships of goal orientation, metacognitive activity, and practice strategies with learning outcomes and transfer. *Journal of Applied Psychology*, *83*, 218–233.

Gagne, R.M. (1984). Learning outcomes and their effects: Useful categories of human performance. *American Psychologist*, *39*, 377–385.

Glaser, R., and Chi, M.T. (1989). Overview. In M.T. Chi, R. Glaser, and J. Farr (eds.), *The Nature Of Expertise* (pp. xv–xxvii). Hillsdale, NJ: Erlbaum.

Gopher, D., Weil, M., and Bareket, T. (1994). Transfer of skill from a computer game trainer to flight. *Human Factors*, *36*, 387–405.

Kozlowski, S.W.J., Gully, S.M., Brown, K.G., Salas, E., Smith, E.M., and Nason, E.R. (2001). Effects of training goals and goal orientation traits on multidimensional training outcomes and performance adaptability. *Organizational Behavior and Human Decision Processes*, *85*, 1–31.

Kraiger, K., Ford, J.K., and Salas, E. (1993). Application of cognitive, skill-based, and affective theories of learning outcomes to new methods of training evaluation. *Journal of Applied Psychology*, *78*, 311–328.

Levy, D.G., Lewis, M.W., Bondanella, J.R., Baisden, M., and Ettedgui, E. (2001). *Exploring the Use of Microworld Models to Train Army Logistics Management Skills*, Santa Monica, CA: RAND, MR-1305-A.

Mathieu, J.E., Heffner, T.S., Goodwin, G.F., Salas, E., and Cannon-Bowers, J.A. (2000). Influence of shared mental models on team process and performance. *Journal of Applied Psychology, 85*, 272–283.

Noe, R.A., and Schmitt, N. (1986). The influence of trainee attitudes on training effectiveness: Test of a model. *Personnel Psychology*, *39*, 497–523.

Noe, R.A., and Wilk, S.L. (1993). Investigation of the factors that influence employees' participation in development activities. *Journal of Applied Psychology*, *78*, 291–302.

Salas, E., and Cannon-Bowers, J.A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, *52*, 471–499.

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

Slife, B.D., and Weaver, C.A., III (1992). Depression, cognitive skill, and metacognitive skills in problem solving. *Cognition and Emotion*, *6*, 1–22.

Stevens, C.K., and Gist, M.E. (1997). Effects of self-efficacy and goal-orientation training on negotiation skill maintenance: What are the mechanisms? *Personnel Psychology*, *50*, 955–978.

Winkler, J.D., Leonard, H.A., and Shanley, M.G. (2001). *Army Distance Learning and Personnel Readiness*, Santa Monica, CA: RAND, MR-1389-A.

Wisher, R.A., and Curnow, C.K. (1998). Techniques for evaluating distance learning events. *Proceedings of the 14th Annual Conference on Distance Teaching and Learning*, Madison, Wisconsin, 461–466.

Wisher, R.A., Curnow, C.K., and Seidel, R.J. (2001). Knowledge Retention as a Latent Outcome Measure in Distance Learning. *American Journal of Distance Education*, *15*, 20–35.