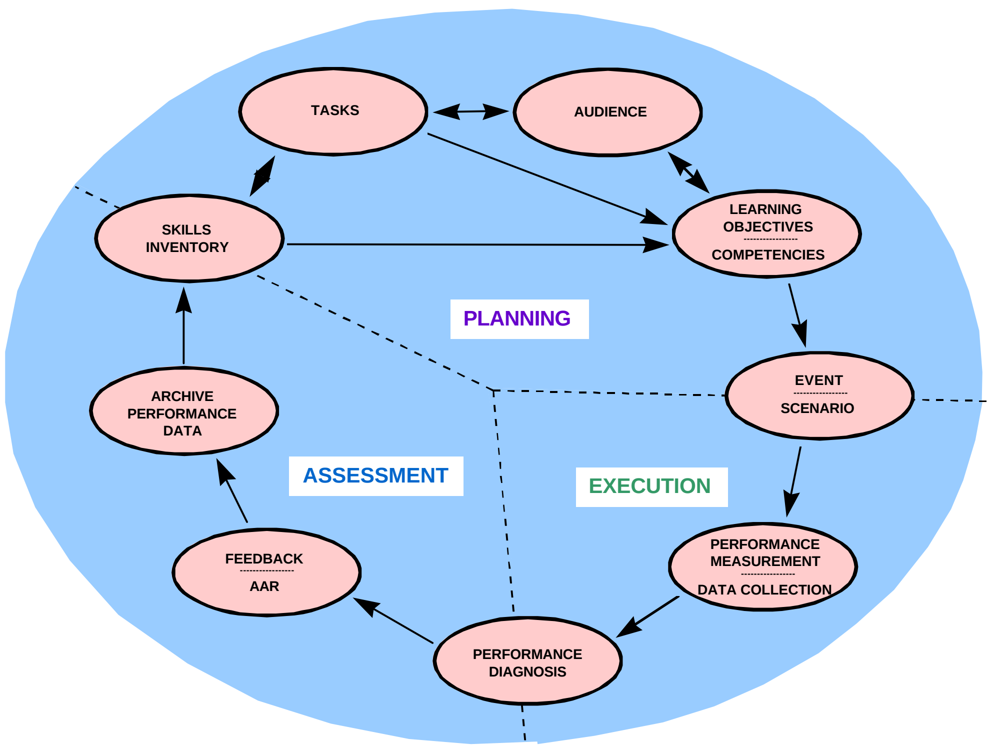
JSIMS
LEARNING METHODOLOGY
REFERENCE DOCUMENT:
A GUIDE FOR
SYSTEM DESIGNERS AND DEVELOPERS



August 1999

Approved for public release; distribution is unlimited.

JSIMS LEARNING METHOLOGY WORKING GROUP

PREFACE

The Learning Methodology Working Group (LMWG) was chartered by OPNAV N7 to assure that the Joint Simulation System (JSIMS) provides an effective learning environment for its users. The initial work of this group was sponsored by NAVSEA PMS430. The Learning Methodology may be thought of as a set of system design attributes, functions, and features necessary for effective learning. The set was derived from behavioral science theory about how people learn and is as close to a *technology of learning* as is available within today's state of the art. Formal acknowledgment of LM in the JSIMS development is necessary because, historically and in the specific case of JSIMS, development priorities tend to be weighted toward engineering rather than effective learning. The teams that oversee system development have different skill sets, perspectives, and priorities than the trainers, educators, and educational psychologists who are most expert at human learning. The LM presented in this document is intended to help bridge this gap by providing system designers and developers with background, guidelines, and practical tools to help them make JSIMS an effective training environment as well as a successful engineering development.

The views expressed in this guide are those of the authors, are not official, and do not necessarily reflect an official policy position of the Services, the Department of Defense, or the U.S. Government.

LMWG COMPOSITION

The members of the LMWG are listed with their various affiliations in Appendix A. In most cases, involvement with the LMWG was for a relatively short term, as members made contributions or represented the interests of their sponsoring organizations before moving on. With very few exceptions, participation in the LMWG was a collateral duty while the member carried on a full-time job in his or her parent organization. LMWG members participated in meetings and conferences; provided documents, information, and written products (such as this guide); offered opinions and critiques; and in other ways supported the LMWG mission. Several authors contributed to the present guide as it evolved over several different drafts. In addition, many of the ideas presented in the guide were drawn directly from their research or work in progress. In alphabetical order, primary contributors to this guide were Jim Brewer (NSWC), Bob Bolling (MITRE), Mike Burke (Novonics), Linda Fenty, Dave Grieve (Sonalysts), Randy Oser (NAWC-TSD), Dan Patton (DSR), Jerry Post (NSC/Cubic), and Henry Simpson (DMDC).

HOW TO USE THIS GUIDE

Purpose

This guide is intended to familiarize JSIMS designers and developers with the basic principles of human learning and to provide guidance concerning the design and systematic evaluation of the JSIMS learning environment to assure that it will be effective for training.

Suggested Reading Strategy

This guide consists of four chapters: 1 (*Introduction*), 2 (*Basic Guidelines for Building an Effective Learning Environment*), 3 (*Building and Evaluating the Learning Environment*), and 4 (*Evaluating the Training System*). Start by reading Chapter 1 to find out what is in this guide, how it is organized, and how the material might be used. Read Chapter 2 to get a sense of how people learn and how this should be reflected in the design and delivery of training in a training simulation. Read Appendix C (Learning Theories) if you are interested in a more detailed discussion of this subject. Chapter 3 is probably the most important chapter in this guide, so please read it with care. It provides background and design tools to help build an effective training system. Chapter 4 explains the role of training evaluation throughout training system development.

What Do You Think?

Although this guide was developed for JSIMS, it is intended to apply to any large-scale simulation. Whether or not it will do this, or do it well, remains an open question. As the guide is still a work in progress, in need of further refinement, regard it as *suggestive* rather than *prescriptive*. If you have reactions to this guide—criticisms, things you like or find particularly useful, suggestions for changes or improvements, etc.—make them known to the LMWG. LMWG lead and POC for questions or comments relating to this guide is Dave Grieve (407-380-5212) dgrieve@sonalysts.com.

EXECUTIVE SUMMARY

Problem and Issues

The development and use of large-scale training simulations such as JSIMS is a recent phenomenon. Such simulations have only come into widespread use in the last decade or so as new simulation and networking technologies have permitted their implementation. They are more complex and costly than the training devices, simulators, and simulations of previous generations by at least an order of magnitude. The risk involved in their development is far from trivial, and the military Services and the Department of Defense face formidable challenges in successfully designing, developing, and effectively using these systems. Historically speaking, and in the specific case of JSIMS, development priorities tend to be weighted toward engineering rather than effective learning. The teams that oversee system development have different skill sets, perspectives, and priorities than the trainers, educators, and educational psychologists who are most expert at human learning. The Learning Methodology presented in this document is intended to help bridge this gap by providing system designers and developers with background, guidelines, and practical tools to help them make JSIMS an effective training environment as well as a successful engineering development.

The LMWG consists of user representatives and subject matter experts (SMEs) in LM. Part of the working group's mission is to ensure that LM ideas are incorporated during JSIMS development, from requirement definition through system delivery. Toward this end, the LMWG participates in activities of JSIMS Integrated Product Teams and working groups. The LMWG offers its LM resources by attending meetings, participating in product peer reviews, and providing LM briefings and white papers to Alliance members. In addition, this guide documents many of the ideas that the LMWG endorses and should facilitate understanding of LM by others.

Objectives

Objectives of this guide are to:

- Familiarize readers with how people learn and how this translates into guidelines for the design and use of simulators to deliver effective training
- Present a conceptual model of an effective JSIMS learning environment and practical tools to help build that environment
- Describe the role of training effectiveness evaluation before, during, and after system development

Evaluation Guidance

This guide is designed to familiarize readers with the basic principles of human learning and to provide guidance concerning the design and systematic evaluation of the JSIMS learning environment. Chapter 1 (*Introduction*) provides an overview and introduction to the guide. Chapter 2 (*Basic Guidelines for Building an Effective Learning Environment*) introduces ten very important principles of human learning based on learning theory. Chapter 3 (*Building and*

Evaluating the Learning Environment) develops conceptual models of the learning environment and human learning processes within that environment. It also provides practical tools in the form of checklists to help designers and developers evaluate the design of a training system before it is implemented. Chapter 4 (*Evaluating the Training System*) deals with training effectiveness evaluation as a training system evolves. It argues that evaluation is important *before*, *during*, and *after* system development. Appendix A (*Learning Methodology Working Group Participants*) lists all current and past members of the LMWG, Appendix B (*Acronyms*) spells out the acronyms used in this guide, and Appendix C (*Learning Theories*) provides a brief introduction to the theories underlying the concepts presented in this guide.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

TABLE OF CHECKLISTS

CHAPTER 1

INTRODUCTION

1.1 Overview

This guide contains the usual front matter (title page, preface, executive summary, table of contents). Guide content is in Chapters 2 through 4, each of which contains several parts. Chapter 1 (*Introduction*) describes the problem and issues, objectives, and shows where the guide addresses each of its objectives.

Chapter 2 (*Basic Guidelines for Building an Effective Learning Environment*) introduces ten very important principles of human learning based on learning theory. The principles are based upon accepted theories about how people learn. Over the years, several learning theories have been examined, evaluated, demonstrated, and/or applied in a variety of operational and laboratory settings connected with military training. There are differences among the theories but many share important common elements; these common elements have given rise to a shared vision of Learning Methodology across the military services. Chapter 2 is intended to help readers understand how the properties of a learning environment (for example, a training simulator such as the Joint Simulation System [JSIMS]) will influence student learning, for better or worse. This chapter should be of particular value to system designers and developers who have not been exposed in their prior education or experience to concepts of learning theory, educational psychology, or training design.

Chapter 3 (*Building and Evaluating the Learning Environment*) develops conceptual models of the learning environment and human learning processes within that environment. It also provides practical tools in the form of checklists to help designers and developers evaluate the design of a training system before it is implemented. This chapter is both the most difficult and most important chapter in this guide. It is the most difficult because it attempts to bridge between the worlds of the educator/trainer and the designer/developer. With very few and rather unique exceptions, the inhabitants of these two worlds have different professional training and experience, speak somewhat different languages, and evaluate the products of their work using different criteria. It is the most important chapter because it provides system designers/developers the essential guidance needed to make their training systems effective for learning.

Chapter 4 (*Evaluating the Training System*) deals with training effectiveness evaluation as a training system evolves. It argues that evaluation is important *before*, *during*, and *after* system development. The chapter develops a framework for deciding what types of evaluation events to stage throughout development to assure that the training system is effective for training, lays out some basic evaluation principles, and describes some tools a training system needs to include to support evaluation.

Appendix A (*Learning Methodology Working Group Participants*) lists all current and past members of the LMWG and their organizational affiliations, phone numbers, and e-mail addresses.

Appendix B (*Acronyms*) spells out the acronyms used in this guide.

Appendix C (*Learning Theories*) provides a brief introduction to the theories underlying the discussion in Chapter 2, along with a glossary defining some technical terminology.

1.2 Problem and Issues

The development and use of large-scale training simulations such as JSIMS is a recent phenomenon. Such simulations have only come into widespread use in the last decade or so as new simulation and networking technologies have permitted their implementation. They are more complex and costly than the training devices, simulators, and simulations of previous generations by at least an order of magnitude. The risk involved in their development is far from trivial, and the military Services and the Department of Defense face formidable challenges in successfully designing, developing, and effectively using these systems. Their promise is great, else there would be no good reason to build them. However, their designers and developers must overcome many problems to assure that these training systems live up to expectations.

1.2.1 Learning Environment and Training Effectiveness

JSIMS development is governed by many specific functional and performance requirements--primarily relating to engineering and modeling. There has been no overall strategy to ensure that it (1) provides an effective *learning environment* or (2) is *effective for training*. These are two separate but closely related questions. An effective learning environment has the attributes, functions, and features necessary to support effective learning by its users. It is the product of a *successful recipe* that combines the *right ingredients* (attributes, functions, features). The right recipe has the potential to provide effective training, although it is no guarantee; many factors will affect the overall outcome. To assure that this outcome occurs, the system must be evaluated systematically throughout development.

1.2.2 Experience Gap

Few people have experience developing large-scale simulations. Each new development draws together teams of individuals--program managers, system designers, programmers, technical subject matter experts, trainers, military users, and others with specialized skills and knowledge--who for the most part are participating in what is for them is a new enterprise. First the *team is formed*. Next it must *set its destination* (define objectives). After that, it must travel to its destination (design and develop system). During travel, it must assess progress along the way (evaluate). To elaborate a little on this geographic metaphor, much of the territory is uncharted and there are few roadmaps to follow. To get to the destination, the team relies on the skills and creativity of its members and their experience in previous developments, as well as in the documented record. Because the territory is unfamiliar, and the team members have different backgrounds, they may not always agree on the objectives, how to travel to the destination, or how to assess progress during travel.

1.2.3 Forming the Team

Developing a training simulation is an engineering task, primarily one of software engineering. Engineers carry out the design and development task under the supervision of program managers who largely reflect the engineering mindset. Other players in development are operations researchers--typically experts on computer modeling--and trainers, educators, and others with specialized technical skills.

1.2.4 Set Destination (Define Objectives)

The JSIMS Concept of Operations (CONOPS) states that JSIMS will provide "computer-simulated environments…to train, educate, develop doctrine and tactics, formulate and assess operational plans, assess warfighting situations, define operational requirements, and provide operational input to the acquisition process." (p. 1). One of the key goals is *training and education*. The present guide focuses exclusively on this aspect of JSIMS.

While the CONOPS clearly states this high-level goal, team members of different backgrounds, training, and experience may perceive it somewhat differently. Unless team members communicate effectively, they may set different destinations--the engineer focusing on hardware/software, the operations researcher on accurate modeling, and trainers/educators on student learning.

1.2.5 Travel to Destination (Design and Develop System)

It is quite possible for an non-integrated team of designers and developers to create a training system that is a successful engineering development but that does not provide an effective learning environment and that does not train anyone how to do anything. To get to the destination, all the developmental interests--engineering, modeling, education/training--must agree and chart a way to get there. While this is a gross oversimplification of the complex interactions involved in an actual system development, it does illustrate the basic problem.

How does one design and develop an effective learning environment? Here, it is important for the engineering and modeling interests to defer to and rely upon the expertise of professional educators and trainers and the body of scientific research they bring with them.

Those with an engineering mindset are often distressed to discover that the education and training discipline has not reached the highly structured and prescriptive level of engineering. True, there are handbooks on training design but they do not begin to approach the degree of specificity of any number of engineering design handbooks. However, educators and trainers in the Service laboratories and elsewhere have been working to translate education and training research into design guidance for training systems such as JSIMS. The guidance exists in the form of a *Learning Methodology* (LM). LM can be thought of as the synthesis of research results and practical experience into a model of how people learn. LM applies directly to military training. It translates directly into a set of system design attributes, functions, and features necessary for effective learning. The set was derived from behavioral science theory about how people learn and is as close to a *technology of learning* as is available within today's state of the art.

1.2.6 Assess Progress during Travel (Evaluation)

As noted above, it is possible to create a training system that is a successful engineering development but that does not provide an effective learning environment. The only way to be sure the system will successfully train people is to *evaluate* it along the way. Because large-scale simulations such as JSIMS are relatively new, the services do not have much experience evaluating them. The DoD Office of the Inspector General (DoDIG) recently conducted an audit that focused particular attention on shortcomings in evaluation of large-scale simulations. In response, the DoD developed revised policy and guidelines for conducting such evaluations. The evaluative guidelines can be used to decide when and how to test and evaluate during system development.

1.2.7 Enter the LMWG--To Bridge the Understanding Gap

The LMWG consists of user representatives and subject matter experts (SMEs) in LM. Part of the working group's mission is to ensure that LM ideas are incorporated during JSIMS development, from requirement definition through system delivery. Toward this end, the LMWG participates in activities of JSIMS Integrated Product Teams and working groups. The LMWG offers its LM resources by attending meetings, participating in product peer reviews, and providing LM briefings and white papers to Alliance members. In addition, this guide documents many of the ideas that the LMWG endorses and should facilitate understanding of LM by others.

1.3 Objectives

Objectives of this guide are to:

- Familiarize readers with how people learn and how this translates into guidelines for the design and use of simulators to deliver effective training
- Present a conceptual model of an effective JSIMS learning environment and practical tools to help build that environment
- Describe the role of training effectiveness evaluation before, during, and after system development

1.4 Road Map: Where This Guide Addresses Each of Its Objectives

The objectives of this guide promise to familiarize readers with the basic principles of human learning and to provide guidance concerning the design and systematic evaluation of the JSIMS learning environment. This guide contains the resources at the locations indicated opposite each stated objective, below:

- Objective 1: Familiarize readers with how people learn and how this translates into guidelines for the design and use of simulators to deliver effective training--See

Chapter 2 (*Basic Guidelines for Building an Effective Learning Environment*). Refer to Appendix C (*Learning Theories*) for an introduction to the theories underlying the discussion in Chapter 2

- Objective 2: Present a conceptual model of an effective JSIMS learning environment and practical tools to help build that environment--See Chapter 3 (*Building and Evaluating the Learning Environment*).
- Objective 3: Describe the role of training effectiveness evaluation before, during, and after system development--See Chapter 4 (*Evaluating the Training System*).

CHAPTER 2

BASIC GUIDELINES FOR BUILDING AN EFFECTIVE LEARNING ENVIRONMENT

This chapter is intended to familiarize readers with how people learn in terms of ten basic principles derived from learning theory. These guidelines are relevant both to system designers/developers and to the trainers who use simulators to deliver training. The learning theory underlying these principles is elaborated in Appendix C. Theory must be interpreted to derive the concrete guidelines needed for practical design/development. The authors of this guide have attempted to make this translation as transparent as possible. Design Checklist 2-1 is an attempt to provide concrete guidance. The checklist summarizes what is necessary to make simulation-based training systems such as JSIMS effective for training. Practically, it will be difficult to satisfy all of these requirements in all situations. However, designers, developers, and trainers should strive to satisfy them as fully as possible inasmuch as they govern the potential training effectiveness of the training system.

Checklist 2-1. Basic Guidelines for Building an Effective Learning Environment

- Provide meaningful interaction between trainees and the learning environment
- Provide repetition of training events
- Vary exercise conditions on successive training events
- Match exercise conditions to learning objectives
- Collect data non-intrusively
- Collect validated performance metrics on team outcomes and processes
- Adjust exercise conditions to match trainee skills
- Provide meaningful exercise conditions
- Provide timely and relevant feedback to trainees
- Conduct After Action Review

Each of the guidelines is described in greater details in the sections that follow.

2.1 Provide Meaningful Interaction between Trainees and the Learning Environment

For learning to occur, actions performed by trainees must influence the situation in the learning environment. “Interaction” implies a two-way relationship. Trainees influence the situation through their behavior, and they in turn get *feedback* from the environment based on their actions. The core function of a simulation-based training system is to provide this interaction. The term “meaningful” implies that the environment’s reactions to trainee actions must be realistic and related to the learning objectives. For example, if a training program is to instill a particular competency, trainees must be able to discern and reflect upon the tangible effects of acquiring or not acquiring that competency.

2.2 Provide Repetition of Training Events

Repetition is essential for learning to take place. Repetition causes knowledge to transfer to and encode in long-term memory. Training must provide sufficient opportunities for trainees to experience the effects of its actions upon the situation's outcome. One of the advantages of a simulation-based exercise is the ability to repeat events. Implicit in this requirement is the need for successive training events to be appropriately timed. Repeat events frequently enough to permit trainees to reflect upon the situation; do not de-couple cause and effect by too-infrequent practice.

2.3 Vary Exercise Conditions on Successive Training Events

As noted, effective learning happens when trainees experience a sequence of training events. Repeating identical events is not generally an effective strategy, for a number of reasons. First, trainee skills change with each succeeding event. For this reason, the trainer/facilitator may, for example, increase the difficulty level for a succeeding scenario. Variation is particularly important when training highly cognitive skills, such as decision-making. Another reason to vary exercise conditions is to reflect changes in the composition of trainees. If a novice replaces an expert on the team, for example, the difficulty of the training event sequence may need to be adjusted. If new combinations of individuals or teams are aggregated into a higher level organization for a particular event, the exercise conditions should reflect the fact that the organization has not trained together.

2.4 Match Exercise Conditions to Learning Objectives

To ensure that the desired behaviors are reinforced, it is necessary to relate the exercise conditions to the learning objectives. If the exercise is not structured with opportunities for trainees to act in the desired way and to experience the consequences of its actions, learning will not occur effectively. If trainees do not view the conditions as realistic, the training may not be perceived as meaningful, which again works against effective learning.

On the surface, this element and the previous one (the ability to vary exercise conditions) may seem redundant. However, given the ability to vary conditions, it is a jump in complexity to recognize and understand the relationships between learning objectives and exercise conditions. Those relationships comprise the essential ingredient in this element of the learning environment.

2.5 Collect Data Non-Intrusively

Assure that the methods of data collection do not influence the behavior of trainees. Data collection should not cause deviation from expected performance or affect trainee behavior.

2.6 Collect Validated Performance Metrics on Team Outcomes and Processes

Trainee performance should be measured using specific metrics related to learning objectives. The data collected will later enable feedback to trainees to support reflective learning and thereby improve performance. Collect data related to both (1) *outcomes* (i.e., was the right decision made?) and (2) *processes* (i.e., was the decision made right?). Outcome measures provide important information regarding overall performance. Process measures provide information on the underlying tasks, sequences of behaviors, and team dynamics that achieved the given outcome. Excellent examples of these metrics are described in *Bibliography* in works authored by Cannon-Bowers, Dwyer, Fowlkes, Oser, and Salas.

2.7 Adjust Exercise Conditions to Match the Trainee Skills

Adjust exercise conditions to suit trainees. There are a number of different factors to consider. One of these is *difficulty*. That is, the higher the skill level, the greater should be the difficulty of the exercise. Other factors are also important. Trainees will almost always consist of more than one individual, and this means that skills may develop at *different rates for different participants*. Certain skills may require more *repetitions to mastery* than others may, even if the level of difficulty is the same. It may be necessary to change the training strategy with advancing skill level, from a highly structured and controlled environment to one that is less so. Trainee skills comprise a complex set of variables. Exercise conditions must be appropriately varied to train effectively on this complex set.

2.8 Provide Meaningful Exercise Conditions

Trainees must perceive the exercise as meaningful to learn from it. In simulation terms, they must regard the scenario as being realistic. There is a difference between (1) realism in trainee *interactions* with the environment and (2) realism in the *scenario.* Realism in *interactions* depends mainly on the accuracy of the simulation models its fidelity to the real world. The designers and developers build this type of realism into the system. Realism in the *scenario* depends upon how accurately the scenario is written by its author and conducted by the trainer/facilitator. Learning will not occur if the scenario is perceived as unrealistic, even if all the object models are realistic. The closer the scenario matches the trainee mission, the better, and the more lasting the learning will be. The trainer/facilitator needs to consider the expected operational use of the competencies to be trained, and to structure trainee experience opportunities as closely as possible to the expected operational context

2.9 Provide Timely and Relevant Feedback to Trainees

Feedback during training tells trainees how well they are performing. It must be timely so that it is associated with the behavior it reflects. It must be relevant so that it helps trainees to adjust behavior appropriately. The requirement for “timeliness” is fairly straightforward and easy to measure. Achieving “relevance” is more difficult. In training involving teams and cognitive skills, it is often difficult to make the connection between actions and their effects. The effects may not be obvious, may be delayed in time, or there may be cumulative effects that are difficult to associate directly.

2.10 Conduct After Action Review

Effective learning requires that trainees have sufficient opportunities to experience the environment, act upon it, and understand and reflect upon actions taken and the relationships to the outcome of the situation. The role of the trainer/facilitator after the exercise is to facilitate the processes of understanding and reflection. The process of facilitation is quite different from the traditional, pedagogic process of delivering feedback to trainees. Learning is enhanced when trainees can *self-discover* the relationships between actions and outcomes. The contribution of the facilitator is to help the process through guidance, team facilitation techniques, and the furnishing of external reference data (for example, ground truth data). Metrics used must assist this process by providing meaningful information in a systematic and structured manner.

CHAPTER 3

BUILDING AND EVALUATING THE LEARNING ENVIRONMENT

As mentioned in the prefatory guidance, the present chapter is probably the most important one in this guide. It provides background and design tools to help you build an effective training system. Please read with care.

The concepts presented in this chapter are based mainly on the work of research conducted at the Naval Air Warfare Center Training Systems Division in Orlando, Florida. For more in-depth information on the subjects covered in this chapter, refer to *Bibliography*. Key contributors to this body of work (listed alphabetically) are Cannon-Bowers, Dwyer, Fowlkes, Oser, and Salas. (The LMWG POC for questions relating to this work is Randy Oser; see Appendix A)

The chapter consists of three sections:

3.1 *Modeling the Learning Environment* describes a conceptual model of the learning environment in simulation-based military training systems

3.2 *Modeling Learning Processes* describes a conceptual model of the learning processes involved in simulation-based military training systems.

3.3 *Evaluating the Learning Environment* lays out a strategy for evaluating the learning environment in terms of criteria implicit in the models developed in sections 3.1 and 3.2.

3.1 Modeling the Learning Environment

3.1.1 Model

Creating an effective learning environment requires an understanding of human learning and use of an accepted learning model. One learning model is shown in Figure 3-1. This model is based on the concept that the most effective way to develop task and team skills is to provide opportunities for trainees to practice those skills within a contextual environment coupled with effective feedback. One method of feedback is *post-reflective dialogue*. Post-reflective dialogue is defined as an after-action review process wherein trainees (1) reflect (relive) what occurred during the execution phase, (2) explicitly challenge perceptions of what happened, and (3), through honest dialogue, change individual and collective perception. In the model, the contextual environment (Synthetic Battle Space) is based on the task, learning objectives, and trainee skill level. Trainees are immersed in the contextual environment where outcomes are dependent on trainee behaviors.

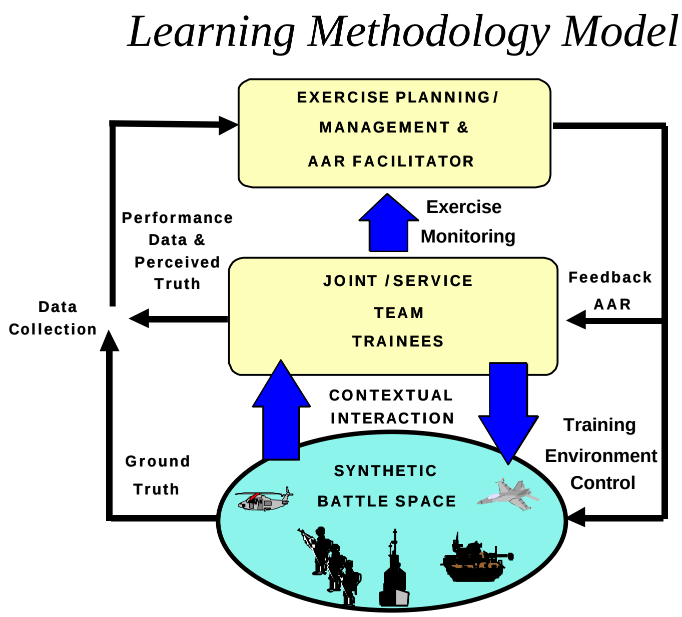


Figure 3-1. A Conceptual Learning Model

3.1.2 Ground Truth, Perceived Truth, and Performance Data

As trainees interact with the environment, data are collected, merged, and stored. Data collected includes ground truth, perceived truth, and performance data. *Ground truth* is the situation and the environment presented to trainees. *Perceived truth* is what trainee sense through interactions with the environment (for example, a C4I display). *Performance data* are information on task outcomes and team processes including data collected automatically and data collected by observers.

3.1.3 After-Action Review

Ideally, during an exercise, trainee performance is monitored, and the complexity of the environment adjusted as appropriate based on skill level or outcomes. After conduct of the interactive exercise, a facilitated, learner-centered After Action Review session is conducted. In this session, trainees (supported by relevant feedback products) re-create what happened during the demonstration period and challenge the results of actions taken. It is during this phase that perceived truth is aligned with ground truth and “discovery learning” takes place. By completing cycles around this model, trainees continue to build proficiency in both task and team skills, which have a direct impact on combat readiness. The team asks the three basic questions:

1. What happened?
2. What should have happened?
3. What do we want to do about it?

3.1.4 Systematic Training and Feedback

Effective learning environments employ systematic, deliberate approaches to ensure skill acquisition and retention by trainees. Certain disciplines need to govern the application of the model to real training situations. For example, efficient learning requires the presence of specific, pre-planned opportunities for participants to demonstrate and receive feedback in targeted competencies (knowledge and the skill to collectively use that knowledge). The introduction of these opportunities must be transparent to maintain realistic trainee performance. Uncontrolled free play, without established learning opportunities and associated feedback, risks wasted resources and failure to achieve objectives. At the same time, scenarios must not be so constrained that the trainee loses interest. A well-designed scenario provides a free play backdrop interlaced with structured unexpected learning objectives.

3.1.5 Data Collection

Learning environments must employ a systematic, coordinated data collection scheme to provide effective feedback on targeted tasks and competencies. The method used for data collection must be non-intrusive.

This learning model supports the types of disciplines required to ensure an effective learning environment, but it does not guarantee that those disciplines will actually govern training. The training system must be designed to support a systematic and disciplined approach to structuring an effective learning environment.

3.2 Modeling Learning Processes

3.2.1 Model

Creating an effective learning environment requires an understanding of human learning processes within the simulation environment. One such model is shown in Figure 3-2. This model reflects the shared conceptual model of the learning environment, and actual experience with simulation based training systems within the military services.
Note that the model divides this process into three parts--planning, execution, and assessment--and that each part consists of several objects, events, or sub-processes that interact and influence one another in predictable ways.

The value of the LM process model is that it expresses LM requirements in terms of a tangible process consistent with military training environments. This is a crucial step in the translation from theory to practice.

Components of the LM process model are described below.

Each component description includes a checklist of specific LM requirements for simulation-based training expressed in terms of the components and their interactions. The lists are included to provide insights into LM requirements and to indicate which software components might be affected by invoking these requirements. Note that the checklists are not intended as a specification for JSIMS or any other specific system but are simply an expression of requirements that support LM concepts.

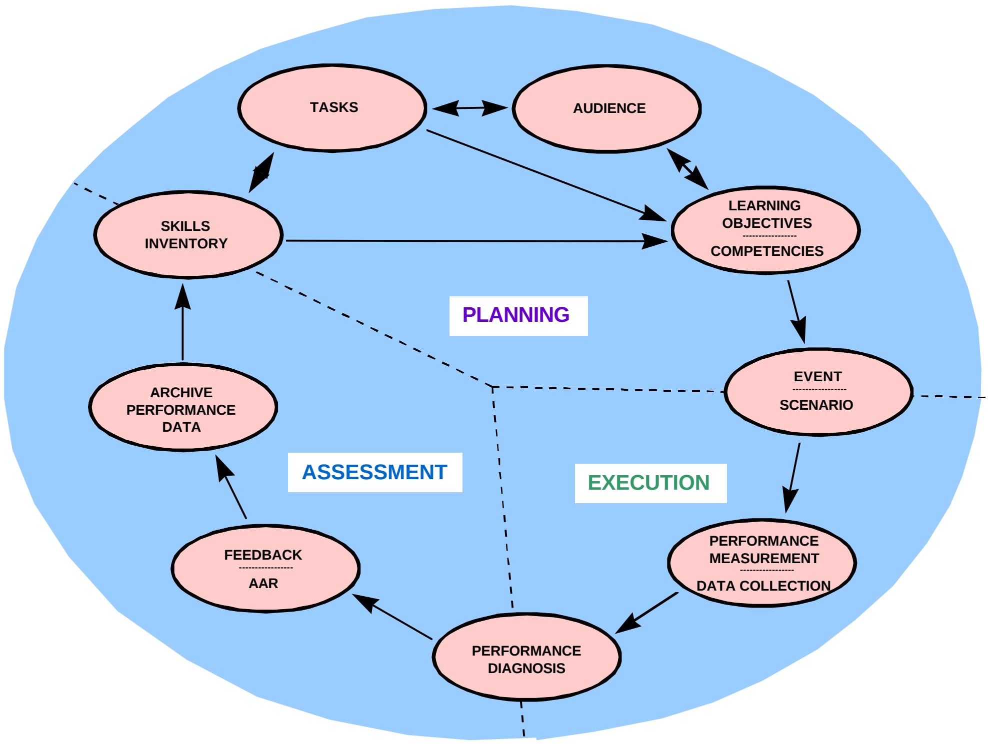


Figure 3-2. LM Process Model

3.2.2 General Requirements

Some general requirements apply to all the tasks or capabilities in a simulation throughout planning, execution, and assessment phases. These requirements address a mixture of efficiency and effectiveness issues. The requirements that fall into this category are shown in Checklist 3-1.

Checklist 3-1. General Requirements

- Shorten clock time or calendar time to perform a function.
- Reduce labor hours to perform a function.
- Reduce skill level requirement to perform a function.
- Reduce the incidence of errors.
- Generate recommendations in support of decisions and selections.
- Maintain schedule of planned events.
- Report scheduled events by member of training audience, organization, task, mission, date, or location.
- Report potential conflicts in scheduling of training audience or training location.

3.2.3 Planning: Trainees and Tasks

Effective training begins with a clear understanding of trainees and training requirements, expressed in terms of mission-essential tasks. One source of task definitions is the Universal Joint Task List (UJTL). The UJTL provides a detailed listing of tasks, conditions, and standards that comprise the missions of a joint military force. There are other useful sources as well, including National Military Strategy, Assigned Missions, Commander's Intent, Joint Doctrine, Joint Tactics, Techniques, and Procedures (JTTP) and service component tactical task lists. The definition of trainees and the required tasks comprise the independent variable of the training event.

Requirements relating to trainees and tasks are shown in Checklist 3-2.

Checklist 3-2. Requirements Relating to Trainees and Tasks

TRAINEES

- Allow trainees to be an individual, an organization, or a collection of individuals and organizations.

TASKS

- Accept task definitions from authoritative sources. *[The system should accept task definitions from existing sources without requiring re-formatting or re-keying.]*
- Allow the trainer/facilitator to modify task definitions.
- Link tasks to performance standards.
- Relate new tasks to skills and standards. *[The user should be able to create a new task definition that is automatically linked to the existing skills inventory and performance standards, based on the new task's relationship to an existing task.]*
- Represent tasks at different levels of aggregation. *[The system should automatically link a given task with the same or higher level tasks at higher organizational echelons.]*
- Allow missions to be defined in terms of tasks. *[The trainer/facilitator should have the option of specifying training requirements in terms of missions, as well as individual tasks.]*
- Allow requirements to be defined in terms of either tasks or missions.
- Suggest and allow editing of task lists for a specified mission and training audience.
- For specified mission and set of actors, provide a skill assessment. *[Skill assessment is the delta between proficiency required for the mission and skills presently held by trainees.]*

3.2.4 Planning: Skills Inventory and Learning Objectives/Competencies

Based upon the trainees and targeted tasks, appropriate learning objectives are identified. These learning objectives represent the "deltas" between the existing skills, as represented in the skill inventory, and the mission-related task requirements.

Part of the discipline of an effective learning environment is the establishment of Measures of Effectiveness (MOEs) and Measures of Performance (MOPs) for each learning objective. MOEs are process measures; they emphasize those actions taken to reach a performance end state. MOPs are *outcome* measures; they focus on the End State achieved. MOEs provide data to answer "Was the decision made right?", while MOPs provide data to answer "Was the right decision made?"

Selected MOEs determine specific data to be collected and the associated feedback products. The data collection infrastructure and tools should generate trends during the exercise, diagnostic performance feedback, external reference products, and assess how well learning objectives were achieved. Data collection across multiple events for a specific learning objective enables one to assess how well an individual or team performed on similar objectives over a range of conditions.

Requirements relating to learning objectives are shown in Checklist 3-3.

Checklist 3-3. Requirements Relating to Learning Objectives

- Report the skills inventory for the specified training audience as it relates to the specified mission or task set.
- Recommend learning objectives based on mission or task requirements, trainees, and skills inventory.

3.2.5 Planning: Event/Scenario

Once the learning objectives are identified, it is necessary to select or create "trigger events" for each learning objective and incorporate these into a scenario. This process reflects the guiding principle that the training must be structured in a disciplined fashion, with a continuously applied understanding of the specific learning objectives. The trigger events create specific opportunities for trainees to practice critical tasks and competencies in a contextual environment, and to experience the consequences of their actions. Typically, a number of events are created for each learning objective that vary in difficulty and occur at different points in an exercise. This not only provides opportunities to reinforce behavior, but also increases confidence in the results.

Once task requirements, learning objectives, trigger events, MOPs, MOEs, and data collection strategies are known, they are amalgamated into a coherent scenario related as closely as possible to the expected mission or operational use of the training being undertaken. Scenarios must permit the trainees to interact in realistic situations that will facilitate transfer of learning from the training environment to the operational situation. Scenarios can use a wide range of constructive, virtual, synthetic, and live resources. Regardless of the specific resources used to create the training environment, the scenario must support the learning objectives, enable the required events to be presented to the participants, and facilitate the collection of data for feedback on the established MOPs, MOEs and other relevant facts and data.

Requirements relating to Events/Scenarios are shown in Checklist 3-4. Note that these requirements also apply to the next topic.

Checklist 3-4. Requirements Relating to Events/Scenarios

- Recommend a sequence of training events leading to achievement of required competencies.
- Sequentially aggregate up to highest specified echelon. *[The sequence of training events may progress logically up through higher levels of aggregation.]*
- Incorporate repetition as required. *[Consider trainee's learning capability in setting the number and timing of repetitions.]*
- Incorporate variation as required. *[This includes variation for randomness, for increasing difficulty, changing skill mixes within the training audience, and progression from more to less structure training situations based on level of expertise.]*
- Set exit criteria for each training phase. *[Not all learning objectives need to be met in order to advance to a subsequent phase.]*
- Recommend training system configuration.
- Allow prompted editing of training system configuration. *[Prompts should suggest potential conflicts between training system configuration and the intended training audience, scenario, and/or data collection strategy.]*
- Recommend and allow editing of the trainer/facilitator staffing and organization to support a specified event.
- Recommend sequence of event preparation steps.
- Recommend building blocks (vignettes) for constructing each training event. *[Vignettes would be linked to specific learning objectives, and they would incorporate trigger events.]*
- Relate the mission requirements to a geographical area. *[It should be possible to define location-specific missions, or to overlay a selected mission on a specified location.]*
- Relate the entity mix (live vs. simulated entities) to the specified learning objectives. *[A particular scenario may be executed with different mixes of live and simulated entities, depending upon who is being trained.]*
- Build a coherent scenario from selected vignettes.
- Develop coherent sequences of scenarios. *[Avoid overlapping benefits and duplication of effort, thereby optimizing use of simulation resources.]*
- Relate each scenario to organization echelon. *[Link the same or similar scenarios used at different levels of aggregation]*.
- Adjust model fidelity to fit the learning objectives, scenario, and trainees.
- Recommend and allow editing of the scenario initial conditions.

3.2.6 Execution: Event/Scenario

After the scenario is generated and tested, it is used to create the synthetic battle space environment for trainees. This represents the transition from Planning to Execution. Exercise management and control of exercise flow are critical aspects of this process. Training participants must be permitted to make their own decisions and to handle the presented situation consistently with doctrine. At the same time, exercise managers must ensure that opportunities are presented which are aligned with the exercise objectives. Critical features of exercise management include:

tracking the occurrence of events and collecting data during those events, ensuring contingency plans are in place to maintain exercise continuity if there is a failure or anomaly, and monitoring scenario scripts to ensure the exercise unfolds in a way that meets exercise objectives (adjusting if necessary).

3.2.7 Execution: Performance Measurement/Data Collection

As trainees perform within the simulated environment, data are collected to support feedback. An important aspect of this function is the fusion of data, from multiple sources, associated with a particular event and associated trainee actions. When an event occurs, relevant ground truth, perceived truth, and performance data must be collected and correlated. The resulting information can be documented, analyzed, and packaged to provide critical feedback.

Requirements relating to Performance Measurement/Data Collection are shown in Checklist 3-5.

Checklist 3-5. Requirements Relating to Performance Measurement/Data Collection

- Set forth performance measurement (data collection) requirements for each scenario.
- Relate performance measurement (PM) requirements to learning objectives for specified phase of training.
- Relate PM requirements to exit criteria for specified phase of training.
- Relate PM requirements to organization echelon. *[Different data will be collected at different levels of team aggregation.]*
- Ensure scenario events and PM requirements are consistent. *[Highlight potential conflicts between the scenario and apparent PM requirements.]*
- Relate MOPs and MOEs to learning objectives.
- Relate MOPs and MOEs to mission tasks.
- Relate MOPs and MOEs to overall mission effectiveness.

3.2.8 Execution: Performance Diagnosis

Performance diagnosis begins in the Execution Phase. Near-real-time assessments permit the trainer/facilitator to monitor performance and ensure that learning objectives are being met. If appropriate, various components of the situation can be adjusted as to difficulty, or new events can be created to target specific skills. If appropriate, a training event can be truncated if it becomes evident that continuation of the event would not be productive.

Requirements relating to performance diagnosis are shown in Checklist 3-6.

3.2.9 Assessment: Feedback/AAR

Through facilitated team dialogue and use of feedback products, trainees can determine (1) what happened, (2) why it happened, and (3) what they could have done to improve the outcome. Feedback products enhance the ability of the team to relive the exercise and provide external reference information that supports non-threatening changes in both individual and team perceptions (learning). The after action review products integrate ground truth data with perceived data and performance measurements. Feedback elements are based on the MOPs and MOEs, which in turn are linked to the trigger events and learning objectives. This approach provides structure and control to training and ensures internal consistency throughout an exercise. Feedback must be timely and in a form that is relevant to the task at hand. It also must be flexible enough to accommodate varied learning styles.

Checklist 3-6. Requirements Relating to Performance Diagnosis

- Base performance diagnosis on MOPs and MOEs.
- Report the effects of observed MOPs and MOEs on task and mission effectiveness.
- Extrapolate performance trends during the exercise to projected achievement of learning objectives.
- Report if difficulty level appears too high or too low for all or part of the trainees. *[Recommend adjustments to the scenario if appropriate.]*
- Assess and report on learning effectiveness of a given training event or sequence of events. *[Learning effectiveness criteria need to be defined.]*

Requirements relating to Feedback/AAR are shown in Checklist 3-7.

Checklist 3-7. Requirements Relating to Feedback/AAR

- Allow trainer/facilitator to customize AAR products.

3.2.10 Assessment: Archive Performance Data

Following the completion of the exercise, appropriate data are stored and archived in a manner that supports the development of lessons learned. Data collected across exercises can facilitate the development of normative databases that would indicate problem areas and may suggest new instructional strategies.

Requirements relating to Archive Performance Data are shown in Checklist 3-8.

Checklist 3-8. Requirements Relating to Archive Performance Data

- Analyze and report trends across successive events, with regard to specified set of competencies and trainees.
- Allow retrieval, fusion and analysis of archived performance data by selected individual or organization. *[Data should be available for a selected organization, regardless of the level of aggregation of the training event in which that organization participated.]*
- Recommend type, timing, and echelon of a subsequent training event. *[Recommend an adjustment to a planned interval between events, if observed performance so warrants.]*

3.2.11 Assessment: Skills Inventory

Archived data includes a skills inventory database. The skills inventory is updated each time an exercise is conducted. Updated skills inventories are then used in the planning of subsequent exercises. The skills inventory thus represents the transition from one cycle to the next in the LM process model and provides the baseline for follow-on exercise planning process.

3.3 Evaluating the Learning Environment

The LMWG Process model (Figure 3-2) is essentially a theory about how people learn within large-scale simulations. The model represents the operations of the simulation in terms of planning, execution, and assessment phases and, as such, offers a rough standard against which a new simulation design can be compared. The LMWG developed the process model as a way to define requirements for an effective learning environment. This model can serve as a useful analysis, design, and evaluation tool.

3.3.1 Gross Analysis (Count Limbs)

Among other things, the process model identifies (1) phases, (2) functions performed in the simulation and (3) their relationships in time. These factors define a simulation at a very gross level in terms of fundamental evaluation criteria:

- Are the phases present?
- Are the functions present?
- Do the phases/functions follow the sequence in the model?

Analyses to answer these questions can be performed with paper and pencil. (In some cases, one may need to call on subject-matter expert [SME] opinion.) Create checklists. Are all the phases there? The functions? Are they performed in the same sequence as in the model? For negative answers to these questions, determine the reasons why. Are the answers reasonable, or would it be better to tailor the design to the model?

3.3.2 Functional Analysis (Fingers and Toes)

Functional analysis focuses on the content of the functions in the system. These factors define a simulator at a somewhat more refined level:

- How are functions defined?
- What is sequence of events in each function?
- How (hardware and software) is function implemented?

Analyses to answer these questions are an extension of those performed above using analytical and possibly SME-based evaluations. For the first two criteria, compare answers for the model and simulation and resolve discrepancies. The third factor (hardware and software) opens new doors. If actual design information is available, some obvious training effectiveness questions offer themselves:

- What design alternatives are being considered?
- Which is the best?
- How can the design be optimized?

How to address these questions depends upon what is available to evaluate and this depends, to a degree, on how far system development has progressed.

3.3.3 Requirements Analysis

Based on the process model, the LMWG generated a list of *requirements* to support an effective learning environment “for a hypothetical simulation based training system.” The authors caution that the list is not a “specification of JSIMS or any other specific system, nor is it the result of an engineering analysis of JSIMS requirements. It is an independent expression of those requirements that support the precepts of [learning].” Sets of requirements are presented in the form of Checklists for Section 3.2. These requirements on these checklists are attributes that the LMWG contends are important in the design of a new simulation. As with attempting to apply the process model, a reasonable first step is to review and adapt the requirements to the design task at hand. Once this is done, they can be used directly, to see if all requirements are met, as well as to suggest evaluation topics for further study and analysis.

CHAPTER 4

EVALUATING THE TRAINING SYSTEM

This chapter addresses the important issue of training effectiveness evaluation. Among other things, it argues that evaluation is important *before*, *during*, and *after* system development. First, it develops a framework for deciding what types of evaluation events to stage throughout development to assure that the training system is effective for training. Second, it lays out some basic evaluation principles. Finally, it describes some of the tools that JSIMS needs to support evaluation.

This chapter is based on *Evaluating Large-Scale Training Simulations* (Simpson, 1999a,b). Readers seeking a more in-depth coverage of the subjects should refer to the source. (The LMWG POC for questions relating to this work is Henry Simpson; see Appendix A)

4.1 Proposed Training Effectiveness Evaluation Framework

4.1.1 Rationale: Why Evaluate?

Evaluations are conducted for a number of different reasons; obvious ones are to:

- Satisfy milestone requirements
- Assure that system performance standards are met
- Demonstrate cost and training effectiveness
- Identify and correct developmental deficiencies
- Identify and correct deficiencies in the management and use of training systems
- Monitor competencies to support planning and execution of training events

All of these are sound reasons to evaluate. From a purely training standpoint, however, the focus shifts to reasons 2, 3, and 4: 2 and 3 because they show that the system works well and justifies its cost in some relatively mature end state; 4, because evaluations can help identify system shortcomings that can be corrected during development.

4.1.2. Evaluation as Total Quality Management

As we tend to think of evaluations as one-shot events that provide definitive results, the least obvious of these reasons to evaluate is 4. Evaluation conducted for this reason suggests that evaluation (1) is not an event but a *process*, (2) is a technique for improving the system being evaluated, and (3) may or may not provide definitive results. In other words, evaluation can be thought of as similar to Total Quality Management (TQM), wherein data pertaining to a process are gathered and analyzed, the process is critiqued, and corrective actions are taken to improve the process. Data pertaining to the revised process are gathered, analyzed, and so forth, in an endless cycle.

4.1.3. Building an Evaluation Framework

In thinking about evaluation, it is useful to start by asking basic questions; for example., *how*, *what*, and *when* should I evaluate? These three questions take on more specific meanings in the context of an actual evaluation:

- How should I evaluate (What evaluation *methods* should I use?)
- What should I evaluate (What *dependent variables* should I measure?)
- When should I evaluate (How should I conduct evaluation events in terms of *time*?)

4.1.3.1 How Should I Evaluate? (Methods). Many different evaluation methods are available. The Defense Manpower Data Center (DMDC) is currently compiling data on several hundred military training system evaluations. Work to date indicates that evaluations tend to use one of four main methods: *experiment, judgment, analysis, or survey*. In general terms, here is how the methods are applied:

- **Experiments:** determine effectiveness based on *observational* data.
- **Judgment-based evaluations:** determine effectiveness based on *human judgments*.
- **Analytical evaluations:** determine effectiveness based on common *analytical techniques* and using common analytical *strategies*.
- **Surveys:** gather data from a sample of a knowledgeable target population and determine effectiveness based on analysis of the collected data.

Each of the methods can, in turn, be performed in several different ways, comprising a set of submethods. Table 4-1 summarizes these four methods (left column), the corresponding submethods (middle column), and their relative frequency of usage (right column) as found in 250 representative evaluations in DMDC's data base. Each of these methods and submethods is described in greater detail with concrete examples in Chapter 3 of Simpson (1999a).

Table 4-1. Frequency of Usage of Common Evaluation Methods and Submethods

METHOD	SUBMETHODS	PERCENT
Experiment (65% of cases)	True experiment	29
	Transfer	9
	Pre-experiment	10
	Test	6
	Quasi-experiment	5
	Ex post facto	6
Judgment (13% of cases)	Users	6
	SMEs	5
	Analysts	2
Analysis (17% of cases)	Evaluate	10
	Compare	4
	Optimize	2
Survey (6% of cases)		6

Each of the methods can, in turn, be performed in several different ways, comprising a set of submethods. The submethods of Experiment are defined mainly based on distinctions made in Campbell and Stanley (1966). The submethods of Judgment are based on respondent category; i.e., the group whose judgments are considered (Users, SMEs [subject-matter experts], or Analysts). The submethods of Analysis are based on differences in the objectives of analysis (Evaluate, Compare, Optimize). The submethods for these two methods were developed iteratively based on analysis of the various cases of their usage in DMDC's database. The distinctions do correspond to differences in usage rather than mere surface characteristics. The Survey method has no submethods. The submethods vary in terms of the cost and difficulty of conducting them and in the authority with which they support conclusions based on their outcomes.

Based on currently available data, the method most commonly used is experiment (65% of cases). Judgment (13%), Analysis (17%), and Survey (6%) are used in far fewer cases. In practice, different methods are sometimes used in combination, although one of the methods is almost invariably primary. Why do these relative numbers differ? Some possible reasons:

- Acquisition regulations generally encourage experiments.
- Among most evaluators and military decision-makers, experiments have greater face validity than other methods.
- Analysis- and Judgment-based evaluations are generally less difficult and costly than experiments and so tend to be used when experiments are not possible.

Note that, to use experiment, a training system *must exist and be functional in some form*. (The system does not necessarily have to be actual, complete, or final. In some cases, it may be possible to use a mockup or simulation to represent the system. Enough of the system must be represented to conduct a meaningful experiment.)

Judgment can be used in a limited way before a system exists (for example, to estimate training potential of a hypothetical design or the perceived need for a system), but usually requires an existing, functional system. On the other hand, analysis can be performed without an existing, functional training system. Analysis tends to be used in two main cases:

- The system is insufficiently developed to conduct an experiment or gather judgment data.
- Evaluation resources are limited.

4.1.3.2 What Should I Evaluate? (Dependent Variables). Many different dependent variables have been used in evaluating large-scale simulations. To date, no set of variables has gained universal acceptance. Thus, it is necessary to start from basics. Simpson (1999a) derives a set of four key variables:

- Reactions
- Learning
- Collective Performance
- Results

One of the simplest and easiest variables to measure is the *reaction* of participants to a particular training experience. This is commonly done with a post-training questionnaire, interview, or, more recently, with a videotaped group discussion akin to an after-action review.

Within the traditional schoolhouse learning paradigm, it is common to evaluate student *learning* based on knowledge and performance test scores. These scores, in turn, may be used to evaluate the training system. In simple terms, the higher the scores, the more effective the training system.

Collective performance--training of groups of people to work together as integrated teams or organizations--is a fundamental a part of military training. Thus, an important criterion for collective training is *how well do collectives perform in the training system?* Definition of evaluation criteria for *collective performance* is still an immature enterprise, more art than science. However, all training—of individuals or collectives—is built upon tasks. Collective training is intended to provide training on collective tasks. At the Joint level, the Universal Joint Task List (UJTL) contains a comprehensive, hierarchical list of the tasks that can be performed by a joint military force; the conditions under which the tasks are performed; and standards of performance. Comparable Service-specific task lists define the relevant collective tasks at the Service level. These task lists essentially define what tasks the Services and Joint forces are expected to be able to perform. They are the logical tasks to use when building scenarios to evaluate collective training.

The operational testing community usually measures performance capabilities in terms of engagement or battle outcomes. There are analogous variables for training systems; that is, what are the *tangible results* during training? (exchange ratio, percent losses by force, shots/kill, etc.) One can measure the performance of the system in terms of achieving its overall objectives while trainees use it. For example, does the simulated Tank Company defeat the simulated enemy; or, do the senior commanders participating in a war game win the war?

In certain circumstances, one might be able to measure *transfer of training* from the system to the real world in terms of combat readiness, field exercise performance, or simulated (or actual) combat.

All of these criteria are of interest, but they are not all of equal significance. While it is useful to gather *reaction* data, they are less important than *collective performance*, which in turn are less important than *tangible results* in the training system. And, since the name of the game is to perform well in the post-training world, *transfer of training* criteria are arguably the most important of all--they may amount to winning or losing a battle.

4.1.3.3 When Should I Evaluate? (Timing). Evaluations are usually considered as one-shot events that answer a question at a particular point in time. This may make sense when evaluating simple things that already exist (for example., an inexpensive training method or medium). It does not make sense when evaluating complex and expensive large-scale training simulations that undergo years of development before becoming operational. For example, early on, evaluations might be conducted to determine whether a prototype design is capable of training on certain tasks. Much later, the total system is put to the test to determine whether it, for example, improves combat readiness.

Table 4-2 illustrates the timing of proposed evaluation events for a large-scale training simulation. Evaluation occurs in four phases, represented by the four right-most columns: *I. Prospective, II. Developmental, III. Milestone,* and *IV. Post-development.* Phases II and III may occur concurrently, if there are several successive milestones during a development. They are separated here to simplify discussion. The left-most column indicates what entries appear in the cells: *When* (timing of events), *Purpose* (why evaluation event is conducted), and *How* (the evaluation method employed). Boldovici and Bessemer (1994) seem to have been the first to advocate a multi-level evaluation strategy such as proposed here. Their recommendations were based on their analysis of the shortcomings of prior SIMNET evaluations. They recommended that several methods be used, as and where needed, to include (1) in-device learning experiments, (2) quasi-transfer experiments, (3) correlational research with archived data, (4) efficient experimental designs, (5) quasi-experimental designs, (6) improved methods for documenting training, (7) analytic evaluations.

Table 4-2. Timing of Hypothetical Evaluation Events for a Large-Scale Training Simulation, Illustrating Purpose and Evaluation Methods by Evaluation Phase

DESCRIP-TION	EVALUATION PHASE			
	I. Prospective	**II. Developmental**	**III. Milestone**	**IV. Post-Development**
When	Before system exists	During system development	At major developmental milestones (builds, IOC, FOC)	After system becomes operational
Purpose	• Estimate perceived need for and training potential of system • Assess/refine design • Define/refine training content • Assure adequate learning environment	• Demonstrate training effectiveness of *functioning subsystems* • Assess/refine design • Estimate user acceptance	• Demonstrate training effectiveness of *total system* • Assess/refine design • Determine user acceptance	• Estimate transfer of training • Determine effects of training on readiness, use of resources, & overall performance • Estimate need to modify training system • Determine acceptance by user population
How	• Analysis • Judgment (SMEs) • Survey	• Judgment (user, SME) • Experiment (pre-, test, quasi-)	• Judgment (user, SME) • Experiment (quasi-, true)	• Judgment (user, SME) • Experiment (ex post facto) • Survey

I. Prospective Evaluation Phase. Evaluation actually starts *before* the system exists. Some purposes of evaluation at this phase are:

- Estimate perceived need for and training potential of system
- Define/refine training content
- Assure adequate learning environment

These questions can be addressed using analytical and judgment-based analyses and survey.

II. Developmental Evaluation Phase. Later, during system development, hardware and software capabilities will be built, in stages, and it will become possible to evaluate these fledgling capabilities. Some purposes of evaluation at this phase are:

- Demonstrate training effectiveness of *functioning subsystems*
- Assess/refine design
- Estimate user acceptance

These questions can be addressed based on judgment (user, SME) and simple experiment (functionality tests, user in-device learning experiments).

III. Milestone Evaluation Phase. As development proceeds, certain milestones will be reached during which relatively mature system capabilities are expected to be demonstrated. At these points, and at the end of development, purposes of evaluation are:

- Demonstrate training effectiveness of *total system*
- Assess/refine design
- Determine user acceptance

These questions can be addressed based on judgment (user, SME) and full-scale experiments. True experiments (involving separate experimental and control groups) are not usually an option when evaluating large-scale simulations, as there is seldom a non-experimental condition to use as a control. If such a control condition is possible, a true experiment can be conducted.

IV. Post-Development Evaluation Phase. After the system becomes operational, it is possible to accrue data to ask questions that could not be addressed during development. Some purposes of evaluation at this phase are:

- Estimate transfer of training
- Determine effects of training on readiness, use of resources, and overall performance
- Apply (above) information to estimate need to modify training system

These questions can be addressed based on judgment, survey, and ex post facto experiments to estimate effects of training on readiness and transfer to the operational setting.

When is evaluation complete? Proponents of TQM would argue that evaluation is never over, as the system can always undergo further refinement. Furthermore, in this technologically evolving world, equipment and doctrine changes may stress the ability of training systems to keep up. A more realistic answer is that evaluation ends when the body of accumulated evidence persuades decision makers that the system has successfully met its goals in terms of support of combat readiness, training, and cost reduction, and there are no significant changes in mission, doctrine, weapons systems/equipment, organization, and job design.

4.2 Evaluation Principles

The evaluation framework is based on the following are the principles:

- Essential first steps; define purpose of evaluation and stakeholders. Gain consensus.
- Evaluation is a process, not an isolated event.
- Evaluators should attempt to influence design and development. (Key: well-formed evaluation plans.)

- Evaluate (1) longitudinally (across time) and (2) vertically (across a family of measures) (This provides more useful decision-making information than point measures.)
- Obtain the best data possible, based on (1) state of development and (2) resources. (The worth of the evaluation is a direct function of the quality of its data.)
- Assure that data are valid and reliable.
- Develop learning curves for the collectives using the training system. (These reflect the rate of learning and can provide an indication of system training effectiveness in the absence of a control group.)
- Measure transfer of training to the job. (These provide an ongoing measure of the validity of training and possible need for change.)

4.3 Tools Needed To Support Evaluation

Evaluation events that rely on judgment, analysis, and survey will usually be based on paper and pencil data collection instruments, protocols, etc., and are not expected to generate any special requirements in terms of design.

Experiments conducted during Phases II and III will be based upon a variety of measures reflecting *Reaction*, *Collective Performance*, and *Results.* Assigned evaluators and/or O/Cs may collect some of these measures. However, to facilitate efficient data collection, it is necessary to provide a semi-automated data collection, storage, retrieval, and display capability. (This may very well tie in to whatever *after-action review* system is already in place.)

This topic was covered in greater depth in Chapter 3. Tools need to be designed and developed to support these requirements; for example:

- Interface with O/C tools
- On-the-fly data acquisition
- Short, medium, and long-term data storage
- Retrieval and display for evaluation purposes
- Archival storage

The archival storage capability is particularly important in evaluating large-scale simulations because of the need to accumulate and integrate data over the long term to separate effects of training from confounding variables.

BIBLIOGRAPHY

Boldovici, J.A. & Bessemer, D.W. (1994). *Training research with distributed interactive simulation: Lessons learned from simulation networking*. TR 1006. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA285584)

Campbell, D.T. & Stanley, J.C. (1966). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally. (Also published as Campbell, D.T. & Stanley, J.C. (1963). *Experimental and quasi-experimental design in teaching*. In N.L. Gage (Ed.). Handbook of Research on Teaching. Chicago, IL: Rand-McNally.)

Cannon-Bowers, J.A. & Salas, E. (1997). A framework for developing team performance measures in training. In M.T. Brannick, E. Salas, & C. Prince (Eds.), *Team Performance Assessment and Measurement: Theory, Methods and Applications.* Hillsdale, NJ: Lawrence Erlbaum.

Department of Defense, Office of the Inspector General (1997). *Requirements planning and impact on readiness of training simulators and devices.* Report No. 97-138. Arlington, VA: Author.

Dwyer, D.J., Fowlkes, J.E., Oser, R.L., Salas, E., & Lane, N.E. (1997). Team performance measurement in distributed environments: the TARGETs methodology. In M.T. Brannick, E. Salas, & C. Prince (Eds.), *Team Performance Assessment and Measurement: Theory, Methods and Applications* (pp. 137-153). Hillsdale, NJ: Lawrence Erlbaum.

Dwyer, D. J., Oser, R. L., Salas, E., & Fowlkes, J. E. (1999). Performance measurement in distributed environments: Initial results and implications for training. *Military Psychology,11* (2), 189-215.

Fowlkes, J., Dwyer, D.J., Oser, R.L., & Salas, E. (1997). Event-based approach to training. *Proceedings of the 19th. I/ITSEC Conference.* Albuquerque, NM.

Garlinger, D.K. & Fallesen, J.J. (1988). *Review of command group training measurement methods.* TR 798. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ADA201753)

Joint Warfighting Center (1997). *Joint Simulation System Concept of Operations Version 1.0 (15 August 1977).* Ft. Monroe, VA: Author.

Oser, R. L., Cannon-Bowers, J. A., Dwyer, D. J., & Salas, E. (1997). Establishing a learning environment for JSIMS: Challenges and considerations [CD-ROM]. *Proceedings of the 19th Annual Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL,* 144-153.

Oser, R. L., Dwyer, D. J., & Fowlkes, J. E. (1995). Team performance in multi-service distributed interactive simulation exercises: Initial results [CD-ROM]. *Proceedings of the 17th Interservice/Industry Training Systems and Education Conference, Arlington, VA,* 163-171.

Simpson, H. (1999a). *Evaluating large-scale training simulations, Volume I: Reference Manual.* DMDC TR 99-05. Seaside, CA: Defense Manpower Data Center.

Simpson, H. (1999b). *Evaluating large-scale training simulations, Volume II: User's Manual.* DMDC TR 99-06. Seaside, CA: Defense Manpower Data Center.

APPENDIX A

LEARNING METHODOLOGY WORKING GROUP PARTICIPANTS

Table A-1. Learning Methodology Working Group Participants

Name/Address	Affiliation	Phone	e-mail
Robert Bolling The MITRE Corporation 3504 Lake Lynda Drive, Suite 155 Orlando, FL 32708	MITRE	407-208-5763	rbolling@mitre.org
Mr. James Brewer (4A21) PHD NSWC Port Hueneme, CA 93043	NSWC	805-982-6747	Brewer_James@phdnswc.nswses. navy.mil
Lt Col Tom Burwell Air Force Agency for Modeling & Simulation (AFAMS) AFAMS/PSD 12350 Research Parkway Orlando, FL 32826-3276	AFAMS	407-208-5746	tom.burwell@afams.af.mil
Dr. Janis A. Cannon-Bowers Naval Air Warfare Center Training Systems Div 12350 Research Parkway Orlando, FL 32826-3276	NAWC	407-380-4830	janis_cannon-bowers @ntsc.navy.mil
Jim Clary NSC/Cubic	NSD/Cubic	407-306-5139	jim_clary@csi.com
Dr. Linda Fenty 8900 Elliott's Court Orlando, FL 32836		407-384-6900	linda.fenty@fti.enzian.com
Rich Goolsby JWFC	JWFC	757-726-6120	goolsbyr@jwfc.js.mil
Dave Grieve JSIMSM/Sonalysts	Sonalysts	407-380-5212	dgrieve@sonalysts.com
Joyce Gummel		407-737-8117	jgummel@tacintel.com
Jim Hutchison	NATSIM	703-808-6536	hutchisj@nro.mil
Dr. Carol Johnson DMDC 400 Gigling Road Seaside, CA 93955-6771	DMDC	831-583-2400	johnsoca@osd.pentagon.mil
Michael McClusky Simulator Systems Research Unit Army Research Institute ATTN: PERI-IF12350 Research Parkway Orlando, FL 32826-3276	NAWC	407-380-4990	MccluskeyMR@navair.navy.mil
Bill Newlon JSIMS M/Sonalysts	Sonalysts	757-686-7485	newlon@jwfc.acom.mil

Randy Oser Naval Air Warfare Center Training Systems Div 12350 Research Parkway Orlando, FL 32826-3276	NAWC	407-380-4818	randy_oser@ntsc.navy.mil
Dan Patton Digital System Resources, Inc. 6965 University Blvd. Winter Park, FL 32792	DSR	407-679-1791	dpatton@dsrnet.com
Jerry Post NSC/Cubic	NSC/Cubic	407-679-1210	jpost@dsrnet.com
LCOL John Roche JSIMS JPO	JSIMS JPO	407-384-5578	
Dr. Henry Simpson DMDC 400 Gigling Road Seaside, CA 93955-6771	DMDC	831-583-2400	simpsohk@osd.pentagon.mil
Steve Slosser	JSIMS JPO	407-380-4599	steve_slosser @jsims.mil
CDR Joe Tofalo JWFC	JWFC	757-726-6419	tofaloj@jwfc.js.mil
LCDR Chuck Walker SPAWAR	SPAWAR	619-537-0215	walkerc@spawar.navy.mil

APPENDIX B

ACRONYMS

AFAMS	Air Force Agency for Modeling and Simulation
ARI	U.S. Army Research Institute for the Behavioral and Social Sciences
C4I	Command, Control, Communications, Computers and Intelligence
CONOPS	Concept of Operations
DA	Development Agent
DMDC	Defense Manpower Data Center
DoD	Department of Defense
DoDIG	DoD Office of the Inspector General
DSR	Digital Systems Resources
JPO	Joint Project Office
JSIMS	Joint Simulation System
JTTP	Joint Tactics, Techniques, and Procedures
JWFC	Joint Warfighting Center
LM	Learning Methodology
LMWG	Learning Methodology Working Group
MOE	Measure of Effectiveness
MOP	Measure of Performance
NATSIM	National Simulation System
NAVSEA	Naval Sea Systems Command
NAVSEA PMS 430	NAVSEA Performance Monitoring, Training, and Assessment Program Office
NAWC	Naval Air Warfare Center
NAWC-TSD	Naval Air Warfare Center Training System Division
NSC	National Simulation Center
NSWC	Naval Surface Warfare Center
O/C	Observer/Controller
OPNAV	Office of the Chief of Naval Operations
POC	Point of Contact
SIMNET	Simulator Network
SME	Subject-Matter Expert
SPAWAR	Space and Warfare Command
TQM	Total Quality Management
UJTL	Universal Joint Task List

APPENDIX C

LEARNING THEORIES

This appendix provides a brief overview of three theories of human learning: Behaviorism, Cognitivism, and Constructivism. The appendix also includes a section on learning within an organization and a glossary of terms. The material is covered at a fairly basic, non-academic level. The intent is to provide the reader with an understanding of some of the theory underlying the concepts presented in this guide.

There is no single universally accepted theory of how people learn. Multiple theories continue to evolve. Moreover, there are large individual differences in how people learn in particular situations. Thus, for the foreseeable future, there will undoubtedly be numerous competing theories on learning, as well as numerous variants of those theories to cover specific types of individuals and learning situations. Despite this diversity, the most accepted theories share certain concepts. The following discussion will show this commonality.

C.1 Behaviorism

The main tenet of the behavioralist approach is that environmental factors shape a person's behavior. This approach is concerned with changes in an individual's behavior that occur as a result of learning. Therefore, the behavioralist focuses primarily on the development of skills and abilities, as opposed to knowledge.

Early behavioralists identified two types of conditioning, referred to as *respondent* and *operant* conditioning, which can affect an individual's behavioral response. Respondent conditioning is a process whereby a subject is conditioned to respond to a certain stimulus from the environment. A well-known example of this process is Pavlov's dog, which was trained to salivate when a bell was rung. This training was accomplished by repeatedly ringing a bell just prior to the dog's receiving food.

Operant conditioning is a process whereby the subject's behaviors work on the environment, and feedback is used to reinforce desirable behaviors. Such feedback may be artificial, such as a reward, or it may be a direct result of the subject's behavior, such as the acceleration of an automobile when the gas pedal is pressed. As in the case of respondent conditioning, this training is the result of repeatedly providing the reinforcing feedback when the desired response is elicited. It is important that the time lag between the operant response and the feedback be relatively short, so that the subject will correctly pair the behavior with the feedback. This requirement is known as temporal pairing. It should also be noted that feedback might be either positive or negative, depending upon whether the goal is to reinforce or extinguish a particular behavior.

The behavioralist approach has been shown to work for relatively simple skills, but it is not effective when more complex tasks need to be learned. This approach is particularly ineffective when there is a strong cognitive component involved (such as decision making), or when temporal pairing is not feasible.

C.2 Cognitivism

Cognitive theorists are concerned with the changes in an individual's knowledge that result from experience with a stimulus environment. The cognitive approach is based upon the concept of *schemata*, or mental models, by which individuals organize their perceived environment. During learning, these schematic structures change by the processes of specialization and generalization. Specialization involves the integration of new information and experiences into existing schemata. Generalization is the process of modifying existing schemata or of creating new ones. For these processes to work in a training environment, it is necessary to provide multiple opportunities for the individual to make changes and additions to existing models based on experience with the environment.

Mental models exist in long-term memory. Therefore, in order for training to be effective, learning must transfer from short-term memory to long-term memory. Different theories have arisen as to the means by which this transfer occurs. What is important to realize is that an effective learning environment must facilitate this transfer. An individual's mere recollection of a training event, even in minute detail, does not by itself assure that learning has taken place, because this recollection may involve only short-term memory. A training program must incorporate multiple exposures – and the *right kind*s of exposures – to the environment and to feedback from it, in order for this transfer to take place.

According to the cognitive approach, in order to ensure that changes in knowledge occur, the learning must be "meaningful." That is, there must be perceived consequences for integrating new knowledge or for failing to do so.

C.3 Constructivism

The constructivist approach is based on the belief that learning is a self-assembly process. Constructivists suggest that individuals "construct" their understanding of a topic area through two processes: conflict resolution and reflection. Within the constructivist framework, discovery learning (i.e., free play) is preferred over formally structured training. Discovery learning requires the trainee to determine the best way of learning; learning is not externally determined or controlled. The responsibility of the instructor is to structure the learning environment to ensure that there are sufficient opportunities to discover instances of the desired learning objective.

It is believed that discovery learning increases a trainee's motivation to learn and produces richer knowledge stores. However, because complex behaviors can be selected and orchestrated by the trainee, it is possible that the trainee's own goals may deviate from those of the training exercise itself. The result is the potential for loss of control over the exercise on the part of the instructor. An additional problem associated with the constructivist approach is that it can lead to idiosyncratic learning, for two reasons. First, the course and progress of a training session will be determined by what the trainee already knows. Second, only that knowledge which is personally meaningful to the trainee will be integrated into long-term memory. These potential drawbacks have led to the suggestion that discovery learning may not be appropriate for novices within a domain. However, as the individual moves towards becoming an expert, discovery learning may foster the development of a richer representation of the problem space. In this regard, it should be noted that modern technological advances, such as interactive and multi-media computers, laser discs, and the World Wide Web, can provide trainees with the tools to support discovery learning when it is appropriate.

C.4 Learning within an Organization

Several types of learning within an organization are suggested by the framework offered by Peter Senge and Michael Marquardt:

C.4.1 Adaptive, Anticipatory, and Generative Learning

Adaptive learning is learning from experience and reflection; for example:

action—>outcome—>results data—>reflection

Learning can be *single*- or *double*-loop learning. Single-loop focuses on gaining information for stabilizing and maintaining existing systems with the emphasis on error detection and correction. Double-loop, a more in-depth process, involves questioning the system itself for the root cause of the errors OR successes.

Anticipatory learning is the process of gaining knowledge by envisioning and expecting the future.

Generative learning is created from reflection, analysis, or creativity.

C.4.2 Deutero Learning

Deutero learning is technically "learning about learning." It occurs when the organization learns from critical reflection on taken-for-granted assumptions. This type of reflection provides an organization the opportunity to discover what they do (or have done) to either facilitate or hinder learning, to invent new strategies to advance learning with the goal of effecting change in the organizational learning practice.

C.4.3 Action Learning/Action Reflection Learning

Action Learning/Action Reflection Learning involves reflecting on real problems using the following formula:

L (learning) = *P* (existing knowledge) + *Q* (questioning insight)

Action learning provides a well-tested method of accelerating learning. When used as a systematic process, organizational learning increases so that it can more effectively deal with change and so that its people can learn better and more effectively handle difficult situations. Action learning is used to examine a complex/difficult task, to move people to act to change it, and to return the results to the organization for review and learning; people devote quality time and energy as needed to learn how to learn and think critically. As a result the individuals involved in action learning build the skills to meet team and organizational needs. Some principles of action learning are shown in Checklist C-1.

Checklist C-1. Some Principles of Action Learning

- Reflection upon participation in an experience increases learning
- If too much reliance is on "expert opinion/information" individuals do not seek own/new solutions
- We learn critically when able to question assumptions that drive an action
- Accurate feedback from others and results of problem-solving actions increase learning
- Working on unfamiliar problems in unfamiliar settings provides greatest challenges; potential and high probability are greatest for learning
- Mixed groups, nonhierarchical, often are better able to gain new perspectives
- Action learning is most effective when learners are examining the organizational system as a whole

Action learning is intended to induce new thinking by conscious consideration of group content, called "an action learning set." The model is centered on the concept that *setting* (or environment) and *problems* to be considered are an important link to group decisions and the depth of the learning experience. These settings and problems can be categorized as either familiar or unfamiliar, but team learning reaches its fullest potential when both setting and problem are unfamiliar.

Michael Marquardt, in *Building the Learning Organization,* identifies specific characteristics for "new learning" said to be applicable in tailoring military training (Checklist C-2).

Checklist C-2. Characteristics for "New Learning" Said To Be Applicable in Tailoring Military Training

- Learning is performance-based, tied to objectives
- Importance is placed on learning processes (learning how to learn) as much, if not more, than on the content
- As important as coming up with the correct "answers" is the ability to define the learning needs
- Across the organization opportunities are created to develop knowledge, skills, and attitudes
- In part, learning is a product of the activity, context, and culture in which it is developed and used
- People are more willing and able to learn that which they have helped create
- Critical survival skill: the ability to know what one needs to know, and to learn on one's own
- Continuous learning is essential for survival and success
- Learning can be accelerated if facilitators help people think critically
- Learning should both accommodate and challenge different learning style preferences
- Learning is part of work—part of everyone's job description
- Learning involves a cyclical, iterative process of planning, implementing, and reflecting on action

C.5 Glossary

Behavioral psychology: the study of overt human behaviors; emphasis is based on the belief that by studying the relationship between environmental events and behavior would lead to an understanding of why humans do what they do without references to their mental processes. Behaviorists (often called S-R psychologists) view environmental factors in terms of stimuli and resultant behavior in terms of responses.

Cognitive psychology: the study of mental process and of changes in an individual's knowledge that result from experience with a stimulus environment. Cognitivists delve into the internal processes by which an individual deals with the complexity of his environment. They also try to define the resulting cognitive structures that he constructs in his mind: the ways in which he perceives and conceptualizes his physical and social world. An important assumption of cognitive theory is that an individual's behavior is always based on cognition, the act of knowing about the situation in which behavior occurs.

Competency: suitable or sufficient skill or skill level, knowledge, or experience to perform a task. A specific job may require multiple competencies. Or a corporation may have several competencies; i.e. several business bases at which they are equally successfully competitive.

Constructive psychology: Based on the belief that learning is a self-assembly process.

Discovery learning: The learner organizes into final form the material to be learned. Discovery learning is like wrapping your own package, reception learning (or didactic teaching) is like having someone open it for you. Jerome Bruner receives primary credit for encouraging discovery learning, whose advantages are: increment in intellectual potency; emphasis placed on intrinsic rather than extrinsic rewards; students masters the methods of "how" to discover; student is more likely to remember information.

Education: The act or process of imparting or acquiring general knowledge and of developing powers or reasoning and judgment.

Evaluation: While measurement only identifies amount, evaluation lays amounts against criteria so that we may make value judgments about the observed amounts.

Knowledge: Acquaintance with facts and truths or principles and familiarity with a particular subject or branch of learning; gained through study, sight, and/or experience.

Learning: Knowledge acquired by systematic study.

Learning Methodology: A systematic process for creating an efficient and effective learning environment which enables a training audience to develop the competencies necessary to perform their required tasks.

Measurement: The process of using numbers to describe quantity, quality, or frequency according to a set of rules.

Skill: Ability to do something well through, talent, training, or practice; learned performance required to complete a task

Task: A logical and necessary step in the performance of a duty—usually a fairly long and complex procedure.

Task analysis: Identification of the behavioral characteristics of a job requirement.

Test: A systematic procedure for comparing the performance of an individual with a designated standard of performance.

Training: The act of making someone proficient by instruction and practice.