



# The Criterion

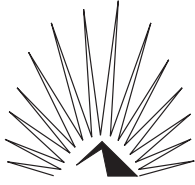
Newsletter of the International Society of Performance Improvement Front Range Chapter

## **Modeling and Simulation: When Do We Really Need It?**

An increasing number of leaders are being pressed by internal staff members and external consultants to invest in modeling and simulation for a wide range of training and education needs ... Via two experiments, we then set out to develop and test a preliminary model that could aid training providers and the leaders they support in (a) determining when M&S is an appropriate training tool and (b) rapidly prototyping the necessary M&S systems.

## **Criterion Q&A with Andrea Moore, M.B.A., CPT and Pat McMahon, M.B.A., CPT**

Andrea Moore & Pat McMahon on challenges in the field, key competencies, and professional resources



# Modeling and Simulation: When Do We Really Need It?

by Pat McMahon, CPT, Andrea Moore, CPT, and Jim Hill, Ed.D., CPT

## Executive Summary

A common perception of mid-level leaders in both the private and public sectors is that training decisions need to be made fast. This tends to result in solution selections made without the benefit of sufficient analysis. The results of these premature decisions are not lost on senior leadership. It is their general experience that the benefit of the typical training intervention is often minimal, while costs are quite high. Despite these general outcomes, an increasing number of leaders are being pressed by internal staff members and external consultants to invest in modeling and simulation (M&S) for a wide range of training and education (T&E) needs.

New M&S are appearing on the T&E horizon more and more frequently. These range from simple board games to virtual reality experiences to specially constructed physical environments staffed with actors. With a trend toward the addition of considerable complexity to M&S, even when a simple tool might do the job, the results are extensive development times, extraordinary development and implementation costs, and a lack of measurable performance results. As Roger Addison, the director of performance improvement for the International Society for Performance Improvement (ISPI) asks, "If M&S is the solution, what is the problem?"

To develop a model that might help answer this question, our organization conducted a deliberate research initiative in partnership with representatives from US defense department organizations (DoD).

It is safe to say that a good chunk of M&S research has been initiated and driven by the US federal government, and it is likely that the DoD is the largest global buyer of M&S. Nevertheless, it was the hypothesis of our DoD partners that *no consistent, rigorous process exists for determining when to use M&S.*

In conjunction with a literature review and interviews with subject matter experts (SME), our research included two experiments. As a result of this study, we confirmed the hypothesis.

Then, via two experiments, we then set out to develop and test a preliminary model that could aid training providers and the leaders they support in (a) determining when M&S is an appropriate training tool and (b) rapidly prototyping the necessary M&S systems. In Experiment 1, baseline M&S selection criteria were used to evaluate a highly complex task to determine if they

provided sufficient decision guidance. Experiment 2 tested an expanded set of decision criteria against five additional high cognitive load (HCL) tasks from a wide range of training courses to determine if the model could be used across a wider spectrum of M&S applications.

## M&S Defined

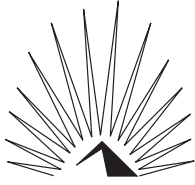
Modeling and simulation is often defined as the use of models, including emulators, prototypes, simulators, and stimulators, either statically or over time, to develop data as a basis for making managerial or technical decisions. In some cases, models are also defined as physical, mathematical, or otherwise logical representations of a system, entity, phenomenon, or process. In most cases though, the terms "model" and "simulation" are used interchangeably.

Within government modeling and simulation manuals, there are descriptions for upwards of 58 types of models including those that are analytic, computational, graphical, and physical. Hence, M&S cut a wide swath. Simple spreadsheet models for "what if?" analyses, maps that represent a real or imaginary terrain, physical simulators for flight training, and highly realistic and complex "first person shooter" games for combat training are all examples of M&S.

## Supporting Performance Objectives

Organizational leaders in both the private and public sectors face similar challenges. In both environments decision makers must be able to envision future situations, conceptualize new systems, and evaluate their performance and manufacturability in a way that carries less risk, is faster, and is less costly. The ability to realize this vision, significantly contributes to organizational success.

Couple these performance needs with shrinking budgets, high turnover, the need for rapid acquisition of expertise, and more demanding "delivery" schedules and it becomes clear that all leaders need a developmental "edge" to ensure their teams can produce the desired results. Time is of the essence, so to address their business needs, leaders are beginning to ask for simple development tools for specific situations, devel-



oped and deployed rapidly for minimal cost. Good, fast, and cheap – the ever-elusive business equation – is a minimum acceptable requirement.

## A Brief Review of the Literature

In the initial approach to the development of M&S selection models, one finds robust data exist supporting the link between performance in the classroom and the field to military readiness (Simpson, 1995). However, in many complex cognitive tasks, it is often challenging to develop definitive performance standards. Simpson goes on to recommend that organizations use expert judgment to capture intangibles such as leadership, motivation, and morale to assess training effectiveness, and that any method of training effectiveness evaluation include some type of SME involvement. In standard operational environments, that may result in a human resource constraint that is untenable or, at least, not cost effective.

In regard to transfer of training, much is said about the requirement for matching media requirements and the training environment. In general, it is agreed that maximum transfer occurs when the training environment and the performance environment (the stimulus and response to performance) are as close to identical as possible.

Cost-benefit considerations are important decision criteria in any performance improvement endeavor. However, with regard to the Department of Defense mandatory cost analysis and training effectiveness evaluations, most DOD instructions, manuals, and directives

provide room for a relatively high financial limit on projects where these techniques are required (Department of Defense, 1991). To aid in the effectiveness of training effectiveness analysis, the US Army developed a handbook for planning, conducting, and writing these events. Unfortunately, it was never published.

In brief, the current state of training development and measurement is one in which senior leaders understand the performance payoff associated with well-developed training. Via doctrine and other formal guidance they have also set the stage for considerable investments to ensure the results of T&E are effectively measured. Yet, the one issue that remains a constant, according to SME interviews, is that more often than not, multiple million dollar technology decisions are being made without the benefit of solid front-end analysis or backend evaluation plans.

## Good M&S Determinations = Many Potential Benefits

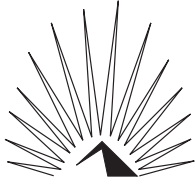
Relevant data regarding M&S needs provides potential benefit for multiple constituents. Table 1 outlines some of the more obvious beneficiaries and the value meaningful M&S decisions could have for them.

## Sexiness versus Performance

There is little question that M&S offers exciting, interactive, and aesthetically pleasing technologies. However,

USER DOMAIN	POTENTIAL USE
Business units, government departments, and their supporting R&D teams	Codify and standardize the process for determining when M&S is an appropriate method for training a specific task.
Training and education organizations	A rigorous, repeatable process for preparing requirements documents for course development for which M&S is a potential training method.  Selection criteria to determine if M&S is an appropriate method for training tasks in the existing curriculum and in future course offerings.
Finance and procurement organizations	Determine if appropriate task analysis, technical justification, and return on investment analysis have been completed prior to contracting for development of courses involving M&S.

**Table 1: Potential Uses of Report Information**



as with certain other technologies, implementation decisions are often made on the basis of newness, graphic detail, and general sexiness – rather than performance impact. That results in other, critical, factors getting ignored. For example, the type of information being imparted to the learner and task criticality are essential in making a solid M&S determination.

To achieve the full benefit of M&S technologies – and to ensure they are applied appropriately in training environments – cost, organizational outcomes, and desired individual performance must first be analyzed (Simpson, 1995). Interviews with SMEs indicated that time, logistics footprint, and personnel restrictions are also significant drivers but rarely considered.

Currently, the typical focus of analysis appears to be on defining the simulation’s functional requirements and physical realism. However, this emphasis on the behavioral and visual aspects of the scenarios often leads to decision-making and cognitive support requirements being downplayed. Regardless of the fact that there are hundreds of simulations on the shelves of most large organizations, it remains unclear whether they have truly contributed to the improvement of individual and team performance.

While there have been many articles, papers, research, and models developed around simulation design (Sticha et al, 1988), cost optimization of simulations (Simpson, 1995), and the overall use of commercial simulation programs in the military (Morris et al, 2002; Belanich et al, 2004), we could find only two resources that outlined when M&S is an appropriate training medium. Those sources (Clark, 1999, 2002) suggest simulation is a preferred media in instructional circumstances requiring complex decision-making.

## The Need for a Replicable Front-end Analysis

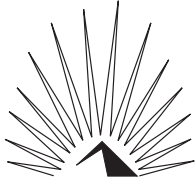
All methods of training system design require effective front-end analysis (Sticha et al, 1990). But it should be of little surprise to learn that rare is the organization employing a standard, consistent, reproducible method of FEA. Moreover, as good as military T&E is – and it is arguably the model for much of the world – Simpson

(1995) adds that it, too, employs no consistent analytical processes, despite spending more on programs and infrastructure annually than any other private or public sector organization. This dramatically heightens the risk of misdiagnosis. Also, in an era of increasing interoperability, jointness, and “transformation,” a lack of cohesive, replicable FEA virtually eliminates the ability to quickly and effectively share related information, reduce redundancy, and maximize the impact of each defense dollar.

Every SME interview reconfirmed that there is no definitive process for determining M&S appropriateness. In addition, multiple SMEs made the following comments:

- The research is subjective and that little documentation exists for selecting M&S or, as an additional concern, for training effectiveness evaluations.
- An important part of effective M&S is the stimulation of decision-making.
- Because of how personnel are globally deployed, with a wide range of environmental situations, M&S is often a poor solution due to the lack of supporting infrastructure.
- The lack of ability to conduct after-action reviews leads to M&S being a poor training tool for field personnel.
- There is a need for well-written course objectives and SME validation before deciding on M&S development and the appropriate level of M&S fidelity.
- Concerns that teaching in M&S environments often involves the use or gaming of the simulation, rather than a focus on learning.
- Concerns that requirements generation processes and cost factors compound the challenges associated with M&S efforts.

A bottom line belief among interviewed SMEs was that the decision to use M&S is often determined prior to any form of task analysis. From their combined perspectives, the question is typically not whether M&S is appropriate, but rather, what type of M&S is desired. This is akin to ordnance decisions being made prior to analysis of the target.



Our researchers took a two-step approach in developing a preliminary M&S decision framework. Initially, complex, high cognitive load (HCL), tasks selected from course curricula were evaluated against the criteria, resulting in each sub-task receiving a weighted score. The decision framework was then tested against five additional complex tasks. The first step was an attempt to validate the criteria. The second was used to determine generalizability.

## M&S Decision Criteria

By dissecting the HCL tasks, the research team was able to identify common factors. The decision framework was derived from these factors and included ten task and two environmental characteristics. A second experiment led to the addition of two more criteria and a much stronger model.

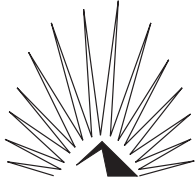
For example, real world tasks, such as those that involve filling out forms or reviewing web pages that involve repetitive information, can often take a long time to complete. Simulations that require completion of a

small sample of the repetitive steps can be much more efficient than real world practice. Additionally, tasks that have built-in wait times (e.g., those involving the completion of a long-running computer program) can also be candidates for M&S. As yet a third example, consider the dependencies on people or organizations that trainees often face in real environment. These might include calling for weather information, requesting special support from an IT organization, or obtaining any other information essential to the completion of the task. In a real-time environment, these external services can be delayed by a myriad of problems such as high demand, shortage of staff, or the need to process higher priority requests.

In the environmental domain, consideration must also be given to the impact that major non-random factors can have on the outcome of certain tasks. Extreme heat, cold, or humidity; hostile local populations; extended periods of light or darkness (related to latitude and time of year); high altitude; and transportation and communication infrastructure are all relevant examples. The initial set of characteristics is shown in Table 2.

Task Characteristics	Environmental Characteristics
<ul style="list-style-type: none"> <li><input type="checkbox"/> Is it hazardous?</li> <li><input type="checkbox"/> Does it include operating large-scale resources or operating in a large geographic area?</li> <li><input type="checkbox"/> Is it costly to execute?</li> <li><input type="checkbox"/> Does it include complex decision-making?</li> <li><input type="checkbox"/> Does it include complex contingency planning?</li> <li><input type="checkbox"/> Does it include problem-solving in an ill-defined domain?</li> <li><input type="checkbox"/> Does it take a long time to execute such that practicing it "real time" is unachievable or unrealistic?</li> <li><input type="checkbox"/> Is there wait time or possible delay?</li> <li><input type="checkbox"/> Does the activity depend upon actions performed by people not included in the trainee group?</li> <li><input type="checkbox"/> Are the task performance measures difficult to achieve in the real environment?</li> </ul>	<ul style="list-style-type: none"> <li><input type="checkbox"/> Is the real environment difficult or undesirable to use or reproduce on demand?</li> <li><input type="checkbox"/> Does the task have aspects of weather, noise, crowds, chaos, or other unpredictable factors that are inherent to the real environment?</li> </ul>

**Table 2. Task and Environmental Characteristics**



Developing decision models is an interesting exercise, but the real question is whether they hold up to scrutiny. The two tests conducted by the development team focused first on the model's mechanics, then on its generalizability.

## Overview of Experiment #1

### Evaluation of sub-tasks

In the first experiment, an evaluation of sub-tasks was needed to complete the decision model. A planning task was selected as the experimental subject due to its complexity (14 sub-tasks). The preliminary M&S selection criteria were matched against the sub-tasks to determine if they provided sufficient guidance. Each sub-task was given a rating based on its contribution to the overall task. A rating of 1 represented a low contribution; a rating of 9, a high contribution. SMEs validated the ratings.

### Criterion importance rating

In addition, each M&S selection criterion was given an importance rating, reflecting the fact that some criteria are more important than others in determining whether M&S is an appropriate training tool. These ratings were based on the initial research for this project and remain consistent for all tasks.

### Selection matrix

The results from the application of the criteria to a task are shown in Figure 1 (on page 7). An "x" was placed in each matrix cell where the characteristics described in the selection criteria were present in a sub-task. Scores were calculated for each sub-task by multiplying its "Sub-Task Importance" rating by the sum of the "Importance of Characteristic" ratings for each of the selection criteria that applied to that sub-task. Scores ranged from 0 to 135. The higher scores indicate sub-tasks for which M&S may be an appropriate training method. Beyond this research, actual use of the criteria would require more clear definitions of the sub-tasks and training objectives, agreement on "Sub-Task Importance" and "Importance of Characteristic" ratings by qualified experts, and a defined method for performing a related training effectiveness evaluation. Only then could a solid M&S recommendation be made, and a rigorous return-on-investment calculation be completed.

### Additional criteria identified

As a result of the initial experiment, we found that two additional task criteria were needed.

- Does the activity involve or rely upon actions performed by people not included in the training group?
- Does the task include complex contingency planning?

### The need for mandatory M&S

Through the research process, it became clear that certain circumstances may lead to M&S being "mandatory." For example, if performing the task in the real environment could result in grave injury or loss of life, initial training for that task ought to include some type of simulation. Criteria that may lead to mandatory M&S decisions include:

- Is the task extremely hazardous or life-threatening?
- Does the task include the need to engage large-scale resources?
- Is the task costly to execute?
- Does the task have aspects of weather, noise, crowds, chaos, or other unpredictable factors that are inherent to the real environment?
- Is the real environment difficult, undesirable, or impossible to use or reproduce on demand?

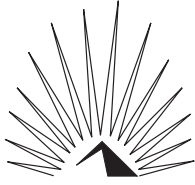
### The need for user's guidance

Without instructions, a model may only be partially effective. The guidelines on page 8 in Table 3 were developed to increase the likelihood that analysts and decision makers employ the model consistently.

## Overview of Experiment #2 – Generalizability

**Testing against additional tasks.** Once the M&S selection criteria model was validated and modifications made, the same process was followed using an additional five tasks. These tasks represented a much broader range of cognitive challenges, and provided a good test of the general applicability of the recommended decision selection criteria. To maintain con-





sistency in the responses throughout the experiment, the detailed task characteristics were referred to when necessary.

**Three specific outcomes.** After making the changes to the decision criteria and then applying the revised criteria set to the remaining five tasks in the model, three outcomes were observed:

**1. A sub-task was encountered that fit the criteria for mandatory simulation.** One of the tasks included a sub-task requiring the conduct of reconnaissance in enemy terrain. It is possible that instructional designers worded this sub-task poorly. However, as written, it is impossible to execute in a live training environment

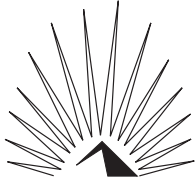
since enemy terrain is necessary. This would need to be simulated in some way.

**2. The criterion, “Are the performance measurements difficult to achieve in the real environment?” was difficult to rate due to a lack of clarity in the course documentation.** This issue was encountered in two tasks. Because the course documentation did not reflect the steps or any type of a process to execute these sub-tasks, it was not possible to judge the difficulty of obtaining performance measures.

**3. Due to the inconsistent use of the terms (e.g., mission analysis) in three of the analyzed tasks, it was difficult to evaluate whether a particular criterion**

USAGE GUIDELINES
1. In order to have confidence in the results of the decision criteria, a rigorous front-end analysis should be conducted so that the tasks and sub-tasks are clearly specified.
2. In general, activities with the highest Sub-Task Scores are the best candidates for considering M&S as an appropriate training method.
3. There are situations in which the case for using M&S is overwhelming, independent of the Sub-Task Score. Examples include: <ul style="list-style-type: none"> <li>• Activities that are extremely hazardous to practice in the real environment (e.g., bomb disposal)</li> <li>• Activities that would involve huge resources to train in the real environment (e.g., large numbers of ships or aircraft)</li> <li>• Activities that involve the consumption of costly resources to practice in the real environment (e.g., the firing of cruise missiles)</li> <li>• Activities that cannot be trained in a real environment (e.g., deployment in an unfriendly nation)</li> </ul> <p>The column headed “Simulation Mandatory” has been provided for designating such activities. It is not necessary to evaluate these activities against each of the detailed criteria, although it may be of interest to compare their Sub-Task Scores to the other sub-tasks.</p>
4. After identifying the top candidate sub-tasks for the application of M&S, it is necessary to perform a Return on Investment calculation, comparing the anticipated cost of developing the M&S approach to the expected savings or mission benefits over conventional training methods. (Reminder: M&S methods can range from simple “What if?” spreadsheet models to a large and costly physical simulators.)
5. Before concluding that M&S training methods are economically or educationally justified, it is necessary to do a thorough Training Effectiveness Evaluation. In the case of proposing a new application for modeling and simulation, a preliminary evaluation can be developed by studying a similar existing application in a comparable domain or using a validated effectiveness forecasting method.

**Table 3: Usage Guidelines**



**was applicable.** This is primarily attributed to the lack of satisfactory detail in the course materials.

**Solid performance analysis and instructional design is required prior to making M&S decisions.** It is generally agreed in the training and education community that there are three elements to good instruction. Each of these elements is dependent on solid instructional design. The primary difficulty in determining M&S requirements, was the lack of satisfactory front-end analysis and instructional materials.

**Training effectiveness evaluation also requires solid analysis.** Excellence in the front-end analysis effort is also needed if a primary T&E goal is a meaningful and actionable training effectiveness evaluation. Desired behaviors and performance outcomes must be defined during the front-end analysis. Ideally, these outcomes will be observable, measurable, and time-bound (Mayberry 2005).

## Conclusions

While the focus of this research was on DoD, the results can be extrapolated to the private sector: There is a need to define a repeatable, consistent, rigorous process and criteria for evaluating the applicability of M&S.

M&S decision criteria must be incorporated as an integral part of determining the appropriate training solution. With a rigorous decision process, leaders and their support functions can be assured that the decision to implement M&S is well thought-out and likely to lead to measurable increases in individual and organizational performance.

## References

Belanich, J., Orvis, K. L., & Sibley, D. E. (2004). *Instructional characteristics and motivational features of a PC-based game*. U. S. Army Research Institute for the Behavioral and Social Sciences Research Report 1822. April.

Clark, R. C. (1999). *Developing technical training*. Silver Spring, MD: International Society of Performance Improvement.

Clark, R. C. & Mayer, R. E. (2003). *E-learning and the sci-*

*ence of instruction*. John Wiley & Sons Inc.

Clark, R. E., & Estes, F. (2002). *Turning research into results*. Atlanta: CEP Press.

Clark, R. E. (2004). *Design document for a guided experiential learning course*. Work Product for TRADOC. Contract DAAD 19-99-D-0046-0004. April 11.

Department of Defense. (1991). *DoD instruction 5000.2: Defense acquisition management policies and procedures*. Washington, D.C.

Morris, C. S., Singer M. J., & Tarr, R. W. (2002). *Low-cost PC gaming and simulation research*. Doctrinal Survey. U.S. Army Research Institute for the Behavioral and Social Sciences Research Note 2003-03. October.

Simpson, H. (1995). *Cost-effectiveness analysis of training in the Department of Defense*. Defense Manpower Data Center Training and Readiness Evaluation and Analysis Division. June.

Sticha, P. J., Singer, M. J., Blacksten, H. R, Morrison, J. E. & Cross, K. D. (1988). *Research and methods for simulation design: State of the art*. HumRRO Final Report 88-27. Alexandria, VA: Human Resources Research Organization. October.

Sticha, P. J., Buede, D. M., Singer, M. J., Gilligan, E. L., & R. J. Morrison. (1990). *Optimization of simulation-based training systems: Model description, implementation, and evaluation*. United States Army Research Institute for the Behavioral and Social Sciences Technical Report 896. June.

Sticha, P. J., Campbell, R. C., & Knerr, C. M. (2002). *Training concepts for virtual environments: Study report 2002-2005*. United States Army Research Institute for the Behavioral and Social Sciences. April.



## Criterion Q & A with upcoming ISPI workshop presenters Andrea Moore & Pat McMahon

During ISPI-FRC's next presentation, Andrea Moore, M.B.A., CPT & Pat McMahon, M.B.A., CPT will review the approach used and the decision support model developed in researching decision criteria to determine when Modeling and Simulation (M&S) is an appropriate training method.

Andrea Moore and Pat McMahon answer ISPI-FRC presenter questions:

• **What three professional resources (books, journals, articles, listserves, and/or websites) do you recommend we should all be taking advantage of to stay current in human performance improvement?**

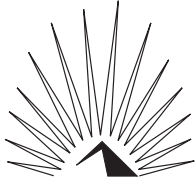
1. Ruth Clark's Building Expertise
2. Zemke and Kremlinger's Figuring Things Out
3. [www.isaga.info/resources.htm](http://www.isaga.info/resources.htm)

• **What do you think is the biggest challenge facing our discipline?**

Companies that want to offer a performance or training "silver bullet". We must always remember that there is no such thing as a quick fix!

• **What are the key competencies professionals in our discipline need to have?**

Number one for is common sense. But I guess this isn't a competency that you can teach. All kidding aside, it all boils down to problem-solving know-how. If you have a feel for how to solve problems, you can pull in the expertise you need for the specifics.



## ISPI-FRC's next presentation 3/14/2006 To Simulate Or Not To Simulate; That Is The Question with Andrea Moore & Pat McMahon

### Event Description

Description of the presentation:

This session will review the approach used and the decision support model developed in researching decision criteria to determine when Modeling and Simulation (M&S) is an appropriate training method. This research was conducted for an organization concerned about spending training dollars wisely, rather than following the latest training technology fads. As part of this project, a new process was proposed, focusing on solving the performance problem, rather than using training as the typical solution.

At the end of the session, participants will be able to:

Select an appropriate definition for a model and a simulation, depending on the organization's structure, background, culture, and predisposition for technology.

Determine when modeling and simulation is a viable solution for training a sub-task.

Use an automated tool to determine a sub-task's feasibility for modeling and simulation.

Technology is often dazzling, but we can't forget that our role as performance technologists is to choose the right solution for the right reasons. This session will also serve as a reminder to all of us that analysis of a performance problem must come first - especially before money is spent on the tool with the most bells and whistles!

During the lecture portion of the session, participants will be challenged to come up with their own definitions of simulation and share them with the audience. Then, working in pairs, participants will have the opportunity to work through two case studies using a paper-based version of the tool. Time permitting, we will input one person's information in the automated tool. We'll end the session by providing participants with an extensive resource list, both paper and online, of some of the latest information sources on modeling and simulation.

### Event Information

To Simulate Or Not To Simulate; That Is The Question--March 14, 2006

Location: Auraria Campus, St. Cajetan's

Members receive discounted admission to all ISPI-FRC events, visit [www.ispi-frc.org](http://www.ispi-frc.org) to become a member.

### Upcoming Events

April 27th & 28th, 2006

2006 Gala Event and Workshop Harold D. Stolovitch, PhD, CPT. Harold is an experienced workshop leader and a keynote speaker at major conferences and organizations throughout the world. He is a prolific writer with 13 books and over 200 articles and book chapters. Harold is co-editor of the first two editions of the Handbook of Human Performance Technology and co-author of the best-selling, award-winning **Telling Ain't Training, Training Ain't Performance**, Beyond Telling Ain't Training Fieldbook and Beyond Training Ain't Performance Fieldbook (in press).

He has won numerous awards for his contributions to the fields of Instructional and Performance Technology including the 2001 ISPI Distinguished Professional Achievement award and ISPI's highest award, Member for Life.

See [www.ispi-frc.org](http://www.ispi-frc.org) for details on events.