



ARROYO CENTER

THE ARTS
CHILD POLICY
CIVIL JUSTICE
EDUCATION
ENERGY AND ENVIRONMENT
HEALTH AND HEALTH CARE
INTERNATIONAL AFFAIRS
NATIONAL SECURITY
POPULATION AND AGING
PUBLIC SAFETY
SCIENCE AND TECHNOLOGY
SUBSTANCE ABUSE
TERRORISM AND
HOMELAND SECURITY
TRANSPORTATION AND
INFRASTRUCTURE
WORKFORCE AND WORKPLACE

This PDF document was made available from www.rand.org as a public service of the RAND Corporation.

[Jump down to document](#) ▼

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world.

Support RAND

[Purchase this document](#)

[Browse Books & Publications](#)

[Make a charitable contribution](#)

For More Information

Visit RAND at www.rand.org

Explore the [RAND Arroyo Center](#)

View [document details](#)

Limited Electronic Distribution Rights

This document and trademark(s) contained herein are protected by law as indicated in a notice appearing later in this work. This electronic representation of RAND intellectual property is provided for non-commercial use only. Unauthorized posting of RAND PDFs to a non-RAND Web site is prohibited. RAND PDFs are protected under copyright law. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use. For information on reprint and linking permissions, please see [RAND Permissions](#).

This product is part of the RAND Corporation monograph series. RAND monographs present major research findings that address the challenges facing the public and private sectors. All RAND monographs undergo rigorous peer review to ensure high standards for research quality and objectivity.

Improving the Army's Assessment of Interactive Multimedia Instruction Courseware

Susan G. Straus, Michael G. Shanley, Rachel M. Burns,
Anisah Waite, James C. Crowley

Prepared for the United States Army
Approved for public release; distribution unlimited



RAND

ARROYO CENTER

The research described in this report was sponsored by the United States Army under Contract No. W74V8H-06-C-0001.

Library of Congress Cataloging-in-Publication Data

Improving the Army's assessment of interactive multimedia instruction courseware /
Susan G. Straus ... [et al.].

p. cm.

Includes bibliographical references.

ISBN 978-0-8330-4727-4 (pbk. : alk. paper)

1. Military education—United States—Computer-assisted instruction—Evaluation. 2. Distance education—United States—Computer-assisted instruction—Evaluation. 3. The Army Distributed Learning Program (U.S.) 4. Computer-assisted instruction—United States—Evaluation. 5. Interactive multimedia—United States—Evaluation. 6. Curriculum evaluation—United States. I. Straus, Susan G.

U408.3.I48 2009
355.0071'5—dc22

2009026430

The RAND Corporation is a nonprofit research organization providing objective analysis and effective solutions that address the challenges facing the public and private sectors around the world. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

RAND® is a registered trademark.

© Copyright 2009 RAND Corporation

Permission is given to duplicate this document for personal use only, as long as it is unaltered and complete. Copies may not be duplicated for commercial purposes. Unauthorized posting of RAND documents to a non-RAND Web site is prohibited. RAND documents are protected under copyright law. For information on reprint and linking permissions, please visit the RAND permissions page (<http://www.rand.org/publications/permissions.html>).

Published 2009 by the RAND Corporation

1776 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138

1200 South Hayes Street, Arlington, VA 22202-5050

4570 Fifth Avenue, Suite 600, Pittsburgh, PA 15213-2665

RAND URL: <http://www.rand.org>

To order RAND documents or to obtain additional information, contact

Distribution Services: Telephone: (310) 451-7002;

Fax: (310) 451-6915; Email: order@rand.org

Summary

Since 1998, the Army's Training and Doctrine Command (TRADOC) has been engaged in establishing and fielding The Army Distributed Learning Program (TADLP) to enhance and extend traditional methods of learning within the Army's training strategy. Distributed learning (DL) is intended to speed the pace of learning and allow training to take place when and where soldiers need it. The Army has an expansive vision of a greatly increased role for DL over time.

Given this expectation, an important component of TADLP's overall performance is the quality of its courses, which consist primarily of asynchronous interactive multimedia instruction (IMI). An assessment of IMI quality is necessary for strategic planning—to understand TADLP outputs, to manage budgets devoted to increasing quality, and to identify and implement needed improvements to processes that affect quality. Moreover, ensuring and documenting the quality of IMI courseware are important to show the value of this type of instruction, to gain the buy-in of DL stakeholders, and to secure the resources needed to achieve the program's goals.

A comprehensive evaluation of training quality requires several types of measures and methods:

- Measures of outcomes, including reactions (e.g., learner satisfaction), learning (e.g., performance on course tests), behavior (performance on the job or in subsequent training), and results (effects of training on organizational outcomes) (Kirkpatrick, 1959–1960, 1994).

- Courseware evaluations to determine whether courses effectively cover the relevant knowledge, skills, and abilities (KSAs) and to assess technical and instructional design criteria that influence the quality of the learning experience.
- Test evaluation using qualitative and quantitative approaches to ensure that tests cover relevant KSAs and can discriminate among good and poor performers. Good tests are needed to measure learning and to assess whether performance in training transfers to performance on the job.
- Administrative data, which can point to potential negative or positive aspects of course quality. Examples include enrollment rates, completion rates, cost data, cycle time of courseware production, and responsiveness of the training development process to the need for changes in courseware content.

Purpose of This Report

This report (1) reviews current efforts to evaluate the effectiveness of IMI in TADLP, (2) develops and tests an approach to evaluating the quality of Army IMI courseware, and (3) identifies directions for improvement in IMI courseware and in the evaluation of IMI within TADLP. As in other approaches to courseware evaluation, we focus on the quality of the learning experience (see, e.g., ASTD Certification Institute, 2001–2003) in terms of instructional design and technical features of courseware, rather than on the accuracy or comprehensiveness of substantive course content. In addition to devising evaluation criteria, we demonstrate the feasibility of this assessment approach by applying it to lessons from a sample of IMI courses. We also illustrate the kinds of information produced by such an evaluation and demonstrate how that information can be used to identify areas for improvement in TADLP courseware and to monitor quality at the program level.

TADLP Does Not Have a Systematic Quality Assessment Process

Our research revealed that there are no efforts in TADLP to assess course quality at the program level. Although several types of independent training evaluation are performed in the proponent schools, these assessments are not comprehensive or systematic in terms of the range of evaluation measures collected, standardization of measures, or synthesis of results across courses. In short, TADLP does not currently have sufficient data or capabilities for data integration to gauge overall program effectiveness or to evaluate the effect of quality-improvement initiatives.

RAND Developed an Approach to Courseware Evaluation for TADLP IMI

We developed a set of evaluation criteria and used them to evaluate lessons from a sample of ten high-priority courses fielded under TADLP between 2005 and 2007. Courses represented a range of knowledge and skills. We had online access to two of the courses, and the remaining eight were provided on CDs.

Our evaluation criteria were based on existing standards in the training development community from (1) the American Society for Training & Development Certification Institute E-Learning Courseware Certification Standards (ASTD Certification Institute, 2001–2003), (2) the Center for Cognitive Technology at University of Southern California (provided by Richard Clark, February 5, 2007), and (3) a checklist developed at TRADOC (provided by D. E. MacAllister, Instructional Systems Specialist, May 16, 2007). We drew criteria from these sources to yield a set of standards that (1) focused on what we viewed as the most important issues, (2) struck a balance between level of detail and feasibility, and (3) could be evaluated from the available materials. Our criteria comprised three categories that reflect the major themes of our source checklists:

- Technical criteria, such as the ease of launching a course and navigating through it, the accessibility and usability of supplementary instructional materials, and the availability and quality of technical support.
- Production-quality criteria, such as legibility of graphics and text, characteristics of audiovisual materials, and the use of techniques to maintain learner interest.
- Pedagogical criteria, such as the specificity of learning objectives; characteristics of instruction of concepts, processes, and procedures; the number and quality of checks on learning; the number and quality of opportunities for practice; and the frequency and quality of feedback.

In addition to applying these criteria, we evaluated the IMI level in each lesson.¹

Three evaluators, or “raters,” were trained in how to evaluate the courses. After establishing satisfactory interrater reliability on a sample of lessons, the raters went through two or three modules of each course in much the same way a student would. They rated a total of 79 lessons, which comprised 39 percent of the lessons per course, on average, or approximately 200 hours of instruction. The majority of the lessons focused on instruction of concepts, followed by instruction of procedures.

Some Features of Courseware Quality Do Not Meet Expected Standards

Our content evaluation found both strengths and deficiencies in course quality. The focus of our evaluation was quality at the program level,

¹ The Army Training Support Center (ATSC) defines four levels of interactivity (Wardell, 2006) which involve progressively greater degrees of interaction between the learner and the computer. These range from instruction in which the learner is a passive recipient of information to immersive learning exercises in which the learner interacts with lifelike visual and auditory cues based on realistic mission scenarios.

but ratings can also be reported at the course level to identify specific needs for improvement.

Levels of Interactivity

In general, we found the content somewhat “thin” in some areas, as exemplified by pedagogical shortcomings and lower-than-required levels of IMI, or interactivity. Our ratings show that 4 percent of the lessons (often introductory lessons) were the lowest level (Level 1), 76 percent included some Level 2, and 20 percent included some Level 3. Whereas some basic material may not lend itself to Level 3 IMI, instruction of even higher-level content consisted largely of Level 2 IMI.

Technical Features

The technical features of the courses were found to be generally strong. All the courses were easy to navigate, and cues to the learner’s position in the course were easily accessible. However, we identified two technical features needing substantial improvement. First, we could not launch most of the courses provided on CDs without substantial technical assistance. We expect that if Army learners experience similar problems and do not have access to technical support, many of them will give up, which would be reflected in low course completion rates. Second, supplementary instructional resources (i.e., reference materials) were difficult to use. Although most courses had a menu through which the learner could easily access resources such as field manuals (FMs), the concepts on any particular page of instruction were not linked directly to an FM. Therefore, using these resources required substantial effort.

Production-Quality Features

Production-quality features of the courses were strong in many areas. Narration was easy to understand, courses had minimal irrelevant content, and most graphics and text were typically legible. However, some aspects of the audiovisual features need improvement. Courses were rated only moderately effective in the use of animation/video to demonstrate processes and procedures. Narration, while easy to understand, was often very slow and could not always be disabled, which inhibited

the pace of learning. Significant improvement is also needed in using media to engage learners and in eliminating sensory conflicts.

Pedagogical Features

The pedagogical aspects of the courseware are the most important criteria in our quality assessment. Pedagogical strengths of the courses include clear descriptions of lesson objectives, an appropriate order of lessons, clear and comprehensive definitions of concepts, and opportunities for learners to correct their strategies in checks on learning and practical exercises.

However, pedagogy was also the area most in need of improvement. A pervasive problem in many of the courses was a lack of examples from the job or mission environments; this occurred in instruction of both concepts and procedures. Courses also need to do a better job of demonstrating procedures and providing explanations of why procedures work the way they do, so that learners can better understand the concepts and skills taught and can thus be prepared to adapt their behavior in nonstandard situations. Finally, in most of the courses we evaluated, practical exercises did not provide sufficient opportunities to integrate concepts and to practice procedures; there were typically too few opportunities, the exercises did not progress from simple to complex problems, and they did not include both part-task and whole-task practice. In short, the courseware was deficient with respect to two critical aspects of training: effective explanations of procedures, and opportunities for practice.

IMI Courseware Should Be Improved in Several Ways

Our analysis suggests a number of potential improvements to IMI courseware:

- Directly linking course content to relevant information in FMs and other supplemental resources will provide learners with a powerful tool for rapidly deepening or sustaining their knowledge in specific task areas. Allowing learners to replace narration

with text could also increase their pace of proceeding through the material.

- Correctly applied, higher levels of interactivity can support improvements in instruction by providing more powerful techniques to improve the relevance and realism of the courseware and the fidelity of practical exercises, thereby enhancing opportunities for transfer of training. Higher-level IMI also can increase learner engagement.
- More generally, in the context of the current phased approach to structured training (in which a self-paced DL phase can be completed up to a year prior to a residential phase), the Army should consider emphasizing the use of IMI to teach concepts and processes rather than procedures. Whereas it is possible to provide practice opportunities for administrative tasks such as filling out forms or logs in IMI (as some courses did), it is clearly much more difficult to provide opportunities to practice physical performance tasks, such as entering and clearing a building, using hand grenades, immobilizing equipment, or tying an eye splice. The use of IMI may be best reserved for training procedures when (1) the procedures can be realistically practiced within the context of IMI (e.g., using software, completing forms, performing calculations) or with the addition of simple and inexpensive job aids; (2) the learning is not subject to rapid decay or is easily refreshed; (3) DL can serve as a supplement to residential training, e.g., when the IMI can be assigned as “homework” immediately preceding a practical exercise; or (4) exported training can be supported by a high level of instructor-student interaction.

We note that TADLP has implemented changes in processes that may have translated into improvements in the quality of courses being developed today. However, the success of these new efforts has not been documented. Moreover, given the range of these initiatives, it is unlikely that all the deficiencies noted in our sample have been corrected. Research is needed to determine the success of new initiatives aimed at improving quality and to guide the Army toward the most effective use of IMI.

TADLP's Assessment Program for IMI Should Also Be Improved

One of the conclusions from our research is that program-level assessments of courseware quality are needed on an ongoing basis. The method used in this study has a number of strengths and can effectively fill at least a part of that need for TADLP. Our approach provides a systematic method of evaluation using multiple raters and a comprehensive set of criteria based on standards proposed by experts in training development and assessment. The method yields quantifiable data, enabling integration of results across courses, schools, or other units of analysis. Once criteria are developed and raters are trained, lessons can be evaluated relatively efficiently. In short, we believe this evaluation method is practical, can provide the Army with valuable information about courseware quality, and points to directions for needed quality improvements. We recommend adoption of this approach for evaluating TADLP courseware.

Our approach to course evaluation does have several limitations. For example, some of the ratings are subjective; all criteria were weighted equally regardless of impact; the sample size of courses was relatively small; and most courses that were provided on CDs did not include some features of the courseware. These limitations can be effectively addressed in future efforts. Consequently, we recommend that TADLP do the following:

- Reassess and refine the criteria used in this study, particularly to reflect fully operational courseware.
- Where possible, establish objective standards for criteria such as the degree to which lessons provide sufficient examples, checks on learning, and practical exercises.
- Establish weights for the criteria according to their relevance to course objectives and the resource implications of remedying deficiencies.
- Evaluate fully functional versions of courses in order to assess features such as course tests, bookmarks, and “live” technical support.

To implement a program of courseware evaluation, we offer the following steps as elements of one potential approach:

- TADLP would assign responsibility for IMI courseware evaluation to a team of two or three individuals trained in principles of instructional design. This task would constitute a portion of each staff member's general responsibilities.
- The evaluation team would work together to refine the evaluation criteria, with input from proponent school staff, subject matter experts (SMEs), and staff from Army Training Support Center (ATSC).
- Proponent schools would provide access to completed courseware as well as to courses under development.
- The evaluation team would evaluate at least a sample of the courseware and provide the completed checklist to the proponent school and contractor, along with specific suggestions for improvement.
- The evaluation team would create quarterly reports and an annual report that summarizes results at the program level.
- These activities would be supported by a website to enable proponent school staff and contractors to download the checklist and supporting material.

We recommend that at the outset, the program focus on providing feedback to individual schools about their courses and on reporting aggregate IMI quality over time. Once the evaluation program becomes more established, the data might be used more aggressively to improve courseware quality. For example, results could be aggregated by school and by contractor as part of a larger evaluation program that fosters continuous quality improvement. Relevant evaluation criteria could be incorporated into the schools' courseware validation processes to standardize and improve these efforts or into the language of DL delivery orders. However, initiating these steps may require formulating new policies to support them, including policies relating to contracts, funding mechanisms, appropriate incentives, and development processes.

As noted earlier, the type of evaluation presented here is only one aspect of assessing training effectiveness. We recommend that TADLP

pursue a comprehensive approach to assessment of training quality at the program level. In addition, many of the methods and measures used to evaluate training can be facilitated by the use of information technology (IT). For example, a broad-based assessment would include the following components, many of which could be tested on a pilot basis using a sample of DL courses:

- **Learner reactions.** The Army could develop a core set of questions to administer to students across IMI courses, develop an IT platform to enable schools to create/customize surveys, and create automated scoring and reporting capabilities.
- **Learning (pre/post comparisons).** The Army could develop an IT platform to administer course pretests and posttests, along with automated procedures to analyze and report test scores across courses or schools.
- **Learning (knowledge retention).** The Army could administer and score follow-up tests relatively efficiently and inexpensively using IT. For example, the Army could administer follow-up tests to IMI learners via Army Knowledge Online (AKO) after students return to their units or when they register for or attend the residential portion of a course after completing the DL phase.
- **Behavior.** IT could be used to collect and analyze data to assess the association of individuals' performance in DL with performance in subsequent residential training or with ratings of subsequent job performance in the unit.
- **Test evaluation.** IT could be used to administer course tests and conduct statistical analyses of objective test items (i.e., multiple choice or true/false) to provide information such as whether items are at the appropriate level of difficulty and whether the tests discriminate between good and poor performers in the course.
- **Administrative data.** Automated systems can capture data such as enrollment and dropout rates, DL usage, and information pertaining to course development, such as cycle time. These types of metrics can be indicators of course quality and should be monitored to assess progress in meeting other TADLP objectives.

Conclusion

This research indicates that a program-level evaluation of IMI courseware is needed and demonstrates a feasible method that can identify areas for improvement in courseware quality. It also suggests other ways in which the quality of IMI training could be practically evaluated at the program level.

In a larger context, evaluation of IMI quality would be part of a more comprehensive assessment component for TADLP. In addition to evaluating quality, that program would include an examination of learning models, courseware usage and efficiency, cycle time of production, and identification of DL areas with the highest payoff for the Army. We are conducting continuing research related to these subjects. Taken together, these efforts will give the Army a basis for managing continuous improvement in the development and use of IMI and will provide important tools to help meet the goals of TADLP.