



## ARROYO CENTER

- CHILDREN AND FAMILIES
- EDUCATION AND THE ARTS
- ENERGY AND ENVIRONMENT
- HEALTH AND HEALTH CARE
- INFRASTRUCTURE AND TRANSPORTATION
- INTERNATIONAL AFFAIRS
- LAW AND BUSINESS
- NATIONAL SECURITY
- POPULATION AND AGING
- PUBLIC SAFETY
- SCIENCE AND TECHNOLOGY
- TERRORISM AND HOMELAND SECURITY

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis.

This electronic document was made available from [www.rand.org](http://www.rand.org) as a public service of the RAND Corporation.

Skip all front matter: [Jump to Page 1](#) ▼

### Support RAND

[Purchase this document](#)

[Browse Reports & Bookstore](#)

[Make a charitable contribution](#)

### For More Information

Visit RAND at [www.rand.org](http://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 electronic documents to a non-RAND website is prohibited. RAND electronic documents 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.

# New Tools and Metrics for Evaluating Army Distributed Learning

---

Susan G. Straus, Michael G. Shanley,  
Douglas Yeung, Jeff Rothenberg,  
Elizabeth D. Steiner, Kristin J. Leuschner

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

New tools and metrics for evaluating Army distributed learning / Susan G. Straus ... [et al.].

p. cm.

Includes bibliographical references.

ISBN 978-0-8330-5212-4 (pbk. : alk. paper)

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

U408.3.N268 2011

355.5—dc22

2011007452

The RAND Corporation is a nonprofit institution that helps improve policy and decisionmaking through research and analysis. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

**RAND®** is a registered trademark.

© Copyright 2011 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 website 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 2011 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](mailto:order@rand.org)

## Summary

---

Distributed learning (DL) is a key element of the Army's training strategy, and the Training and Doctrine Command (TRADOC) has goals for expanding the future use of DL and changing how DL is developed and delivered. Although some individual proponent schools evaluate aspects of DL, the Army currently does not assess training outcomes of The Army Distance Learning Program (TADLP) as a whole, i.e., across proponent schools. Program-level evaluation can play an essential role in the expansion of DL, in the identification of strategic directions for the DL program, and in ensuring the quality of training. First, careful assessment of new initiatives and pilots will help guide the choice of options that best leverage available funding. Second, evaluation will help the program compete for needed resources to support its expansion. TADLP resources for developing content have declined in recent years relative to other training accounts (Shanley et al., forthcoming); better documentation of the value of DL and its contributions to readiness could be essential to substantiate the case for increased funding. Third, evaluation can help identify specific areas for improvement in existing courseware content and delivery, technical matters, course management, training policy, and other aspects of DL design and implementation.

This report describes a series of studies designed to develop and test new tools and metrics to assess training and to document the impact of Army DL courses at the program level. The project on which this report is based builds on prior RAND research that has evaluated different aspects of Army DL. As a foundation for the current research,

we used a model of training evaluation (Alvarez, Salas, and Garofano, 2004), which provides a broad conceptual framework to guide efforts to assess the quality of training and determine how to improve training programs. Below we list the components of the project:

- First, we developed and implemented two online surveys of students' experiences in DL: a *nongraduate survey*, focused on diagnosing why students fail to complete DL courses, and a *graduate survey*, which assessed reactions to completed courses. The surveys provide direct feedback from the ultimate customer (students), and are an important input for documenting DL's current contributions as well as for identifying the key improvements needed in courses and supporting systems.
- Second, we conducted assessments of *knowledge retention* of DL material and the association of learning and knowledge retention with individual soldier readiness for (i.e., performance in) follow-on resident training. Understanding the degree to which students learn from DL and retain knowledge as well as the extent to which DL influences soldiers' readiness for subsequent training is critical to documenting the value of DL and making improvements in future DL content and delivery.
- Third, we *assessed the capabilities of Army information systems* to facilitate enterprise-wide evaluation of DL outcomes. We approached this by interviewing Army subject matter experts (SMEs) concerned with information technology (IT) integration issues involving learning management systems (LMSs) and the Army Training Information System (ATIS). Assessing training effectiveness at the program level requires methods to collect and synthesize the data efficiently. Thus, a fundamental issue addressed in this study is how IT can be used to collect data to evaluate training and do so in a centralized, standardized way.

In the remainder of this summary, we highlight key findings from these three studies, along with specific recommendations relevant to each study. We conclude with recommendations to guide future Army evaluation efforts.

## Surveys of Students' Experiences in DL

Understanding students' experiences in DL has received limited attention at the program level. For example, while individual proponent schools and centers assess student satisfaction with DL, such efforts are sporadic, and the use of different methods and metrics precludes integration of results (Straus et al., 2009). Prior research also shows that many key DL courses have low graduation rates, as indicated by the Army Training Requirements and Resources System (ATRRS), especially when compared to rates for resident courses (Shanley et al., forthcoming). However, the Army does not know why students do not begin courses after enrolling or why they fail to complete courses once they start.

Consequently, we developed two online surveys to assess student reactions in DL training courses, one focusing on nongraduates and one focusing on graduates. Each survey was pilot-tested in five high-priority Department of the Army (DA)-directed DL courses.<sup>1</sup> The non-graduate survey included 1,058 students (a 23 percent response rate from a larger sample), while the graduate survey included 431 students (a 30 percent response rate).

### **A Large Proportion of “Nongraduations” Were Due to Non-DL-Related Factors**

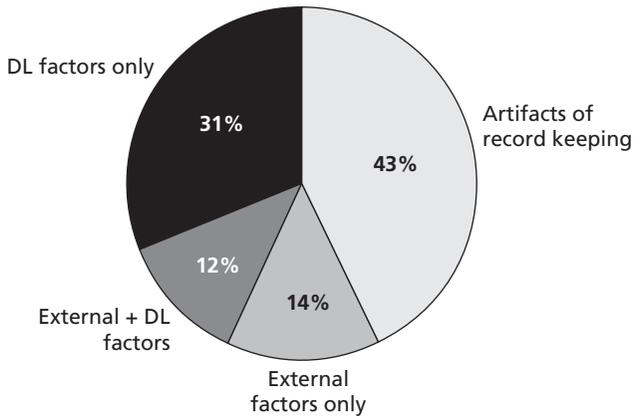
We examined the reasons for nongraduations and categorized the results as shown in Figure S.1.

Results show that in 43 percent of the cases, students who appear as “nongraduates” in ATRRS had, in fact, already graduated or were taking the course for self-development, for which graduation is not required. Thus, in a large number of cases, students' nongraduation status was an artifact of record keeping, and graduation rates are higher than those derived from ATRRS. In addition, in 14 percent of the cases, nongraduation was attributable only to factors external to the DL program, such as mobilization or deployment or changes in occu-

---

<sup>1</sup> Most Army DL is currently part of a phased approach to training. Typically, a DL phase serves as a prerequisite for a resident phase of instruction. Some courses have multiple DL and resident phases.

**Figure S.1**  
**Reasons for Nongraduation (Nongraduate Survey)**



RAND MG1072-S.1

pations. Thus, over half the nongraduations (or cases that appear as such in ATRRS) cannot be attributed to DL courseware or policy.

### **The Main DL-Related Reasons for Nongraduation Were Related to Technical, Support, and Time Issues**

Forty-three percent of nongraduations (DL factors only and external + DL factors in Figure S.1) could be tied to the DL program or its policies, either exclusively or in combination with external factors. The main DL-related reasons for nongraduations were technical problems, a lack of support, and insufficient time to complete coursework.

Two common technical issues for nongraduates involved lack of access to an Internet connection or to a reliable computer. The majority, however, had other technical difficulties. We did not further differentiate among other technical issues in this survey; however, responses to the end-of-course survey described below suggest the kinds of other technical problems that nongraduates may have experienced.

Issues with support centered on insufficient administrative and technical assistance. About 40 percent of these responses constituted cases in which students sought but did not receive help, primarily for technical issues. However, responses from the remaining 60 percent of

these students indicated that simple administrative actions (e.g., follow-up notifications) would address their problems. For example, many who did not graduate were not aware that they had been enrolled in a DL course. This outcome points to issues with the existing policy of automatic enrollment in DL courses, suggesting the need for some follow-up.

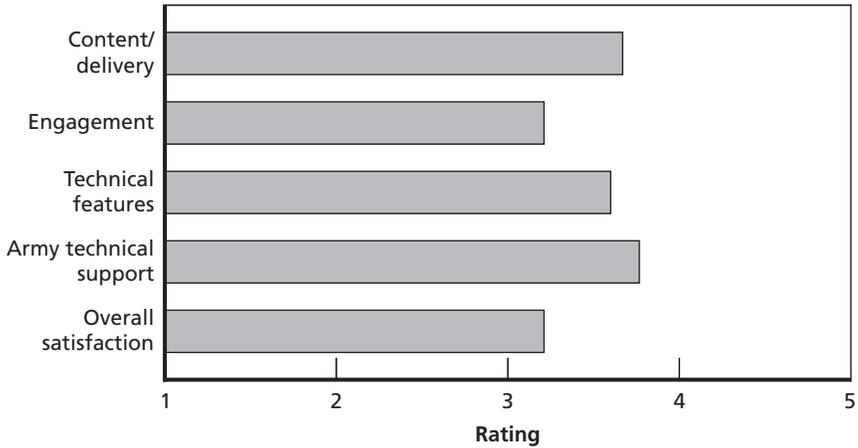
Respondents with “time” issues cited a shortage of time to work on DL. Most of these students were also trying to complete the courses on personal time. Moreover, nongraduates were much more likely than graduates to use personal time to work on DL, suggesting that the amount of personal time used for required DL negatively affected graduation rates. Beyond the effect on graduation rates, this is a larger DL issue because DA policy states that students should be given duty hours (i.e., paid time) to complete required DL training.

Of equal importance are results indicating that some aspects of DL had a minimal effect on nongraduation rates. For example, issues related to DL courseware itself, such as course length or content, played a relatively minor role in explaining nongraduations. Furthermore, comparisons between nongraduates and graduates suggest that failure to complete courses cannot be attributed to an aversion to DL among students. Measures of student learning preferences show that while both nongraduates and graduates tended to prefer classroom learning to DL, the average rating for DL orientation (i.e., a preference for independent, self-paced, computer-based learning) was generally high among all students.

### **DL Graduates Expressed Moderate Satisfaction with Their Experiences in DL Courses, but Responses Also Suggest Areas for Improvement**

Figure S.2 shows average ratings of graduates’ satisfaction with multiple aspects of their experiences in DL courses. Three of the five measures in the figure are based on scales consisting of multiple items. On all five measures, the average rating was greater than 3 on a 5-point scale, suggesting that students were moderately satisfied with all aspects of the course that we assessed. The highest ratings (3.77) went to the technical support provided by Army and proponent school help desks.

**Figure S.2**  
**Graduates' Satisfaction with Aspects of DL**



NOTE: Response options ranged from 1 = strongly disagree to 5 = strongly agree.

RAND MG1072-S.2

Students also gave generally favorable ratings to the quality of course content and delivery. However, ratings of the degree to which courses held the students' interest (engagement) were lower and suggest some need for improvement, especially because engagement was associated more strongly with overall satisfaction than were ratings of other aspects of DL.

Respondents were moderately satisfied with technical features of the courseware. However, 63 percent of respondents reported technical difficulties with the courseware, suggesting some need for improvement. The most common issues pertained to bandwidth or speed (e.g., delays in pages loading; difficulty playing audio or video files), followed by access to courseware (e.g., difficulty launching the course or receiving CD-ROMs). Fewer students reported problems navigating through courseware. Production quality (e.g., difficulty reading the text, excessively slow or fast narration) was the least problematic category of technical difficulties.

## **Issues Raised in Nongraduate and Graduate Surveys Suggest Recommendations for Change**

**The Army needs to capture information about student status in courses more accurately.** Based on findings from the survey of non-graduates, we anticipate that modest changes to ATRRS and to administrative practices could result in significant improvements in DL outcomes:

- We recommend adding a field to ATRRS to document the student's purpose in enrolling in a course (i.e., requirement, self-development, refresher/reachback).
- The Army should also take steps to encourage updating of the ATRRS course graduation field for DL phases.
- Another option is to require that ATRRS have a valid graduation status for the DL phase in order for students to graduate from the resident phase of the course.
- There is also a need to enhance administrative support to ensure that students are aware they are enrolled in DL courses and to offer assistance when progress in the course appears lacking.

**The Army should enforce the policy of paying soldiers for required training.** The Army needs to enforce the policy allowing soldiers to use duty time for required training. One option is the current effort to create an "EDY" or educational duty status that students can use while working on required courses. Another option used by the Army National Guard (ARNG), observed in our studies of DL knowledge retention, is to allow students to come to the schoolhouse one week before the resident course to complete the DL prerequisite on paid time. This approach appeared to increase the likelihood that students would complete the DL phase and that they would do so prior to the start of the resident phase. Other options may also be needed. Absent an acceptable solution, the Army may want to move toward a particular type of blended-learning model (e.g., similar to the ARNG approach described above) wherein students complete the DL content while in residence.

**The Army should seek to improve learner engagement and specific technical features in DL courseware.** Based on the findings of the graduate survey, the Army should seek to increase interaction and to better engage learners in DL training. The Army's current move towards blended learning (which combines different modes of instruction, such as DL and resident training) may ameliorate some of these issues by making use of different media and varied methods of instruction and by providing more opportunities for interaction with instructors and peers.

The nature and frequency of graduates' responses indicate that improvements to technical features of DL courseware should focus on enhancing speed and access. Until bandwidth can be increased, one workaround (suggested by many of the students) is to provide low-bandwidth versions of courses or CD-ROMs for students in deployed locations or in other constrained settings.

**The Army should administer surveys of students' experiences in DL on an ongoing basis.** Our findings indicate that the nongraduate and graduate DL experience surveys are feasible and appropriate for ongoing use at the program level. The scales are psychometrically sound in terms of providing reliable measures, and results show a reasonable degree of variation in responses. From an administrative standpoint, the surveys are not burdensome to complete, and they are straightforward to interpret and score. The surveys can be adapted to address specific goals or topics of interest to TADLP or individual proponent schools.

We recommend that, moving forward, the Army use a common set of indicators on the surveys (with the option for schools to add questions to address local interests), as well as a common platform and software application. Use of both the graduate and nongraduate surveys should be part of routine quality improvement efforts.

## **Knowledge Retention of DL Material in the Phased Approach to Training**

In the next part of this research, we conducted pilot studies to assess knowledge retention in two high-priority courses that use DL in the phased approach to training: Ordnance Mechanical Maintenance Basic Knowledge and Skills Course and the Battle Staff Noncommissioned Officer Course. The first course provided substantive results. The second course did not produce enough data to analyze the association between DL learning and knowledge retention, because some students did not complete the knowledge test (which was voluntary) and others had not completed the DL phase prior to resident training. However, results yielded important lessons for future studies with respect to student participation, DL policy, and Army information systems.

In the Ordnance course, we analyzed the level of knowledge retention following the DL phase; the effect of the lag following the DL phase on knowledge retention; and the associations of learning, retention, and readiness with organizational and student characteristics.

### **Knowledge Retention Was Associated with Time Spent on DL, Lag Time Between DL and the Resident Phase, Job Experience, and Cognitive Ability**

We found, somewhat surprisingly, that the amount of time students worked on the DL phase was not associated with their performance on knowledge tests during DL, and performance in the DL phase was not associated with knowledge retention measured at the beginning of the resident phase. However, there was a positive correlation between time spent on DL and knowledge retention scores, indicating that students who spent more time working on the DL phase retained more knowledge. Greater knowledge retention was also associated with shorter lags between the time that students completed the DL phase and started the resident phase; greater relevant job experience (more civilian experience or longer time in their new military occupational specialty or MOS); and greater cognitive ability (as reflected in higher Armed Forces Qualification Test scores).

## **Recommendations to Address Issues Raised in the Knowledge Retention Assessment**

Results of this analysis suggest directions for change in training policies and procedures to improve DL effectiveness:

- **Students should be encouraged to complete the DL phase of the course in a way that minimizes the time lag between the DL and resident phases.** In addition to changing the policy regarding when students can complete courses, the proponent schools should work to arrange training schedules to enable students to take the DL phase in close chronological proximity to the resident course but without “cramming” DL training into a short time period.
- **Participation in the DL phase may not be necessary for all students, particularly those with relevant experience in the subject area.** Data about relevant student characteristics could be used as a factor in determining whether some students can place out of all or part of the DL phase or be given a streamlined version of the DL course. Pre-tests could be used to accomplish the same goal.
- **DL courses might be more effective if students were supported in planning for sufficient time to complete the course.** We recommend sending students a “welcome” message upon enrollment that provides a timetable to progress through the course as well as contact information for further support. Instructors could also use system logs from the course LMS to identify and contact students who are not making steady progress.

**Other changes could improve the Army’s ability to evaluate knowledge and training performance.** Rather than providing substantive results, the study in the Battle Staff course as well as our experience in the Ordnance course yielded a number of lessons about evaluating knowledge (or skill) retention and training performance:

- **Future studies should be conducted as part of ongoing Army quality improvement efforts** (like the process used in the Ord-

nance course) rather than as a research activity in which participation is voluntary.

- **Enforcing the policy of completing DL prerequisites in advance is also important** in order to study knowledge retention and performance in the phased approach to training.
- **Test evaluation is needed** to ensure that course tests are reliable, valid, and discriminate among high and low performers in the course.
- **Finally, a lack of interoperability among systems that contain relevant data poses a barrier to evaluating learning.** Those challenges may be somewhat less of an issue if such studies are conducted within the Army where student identifying information can be shared more easily.

## **Assessment of the Capabilities of Army Information Systems to Support Enterprise-Wide Evaluation**

We conducted interviews with SMEs in TRADOC Headquarters, selected proponent schools, and Program Manager Distributed Learning Systems (PM DLS) who are concerned with IT integration issues involving LMSs and the Army Training Information System (ATIS). Our interviews with Army SMEs examined both technical and organizational factors that can affect data collection.

### **SMEs See Value in Evaluation Data**

SMEs discussed student-level outcomes, such as performance on course tests, course usage statistics, and student reactions, and course characteristics, such as interactive multimedia instruction (IMI) levels, delivery modes, type of developer, and graduation rates. SMEs generally felt that gathering, sharing, and analyzing these types of data, if done well, could have significant value to schools, course developers, Army Training Support Center (ATSC), and possibly commanders and students. Most respondents felt that sharing data among schools might create a useful exchange of ideas and best practices for producing effective DL. A majority also felt that capturing student-level and course-level

data could provide other benefits, including documenting the value of DL courses, justifying the need for resources, aligning programs of instruction with current doctrine, and providing data needed to better understand the effectiveness of online course development and delivery techniques.

### **However, There Are Numerous Technical Barriers to Data Collection, Sharing, and Analysis**

System and usability problems, such as system crashes, are one source of technical issues that threaten the validity of training evaluation data as well as undermine the training process. However, the predominant technical issue threatening enterprise-wide training evaluation is poor interoperability among information systems. Evaluating training requires finding and querying multiple information systems, most of which are not connected to each other and so require specific expertise or authorization (or both) to access and use. Furthermore, identifying and joining data across systems can be problematic because each system may define and encode data in unique ways. Some systems are connected to each other using one-to-one (“pairwise”) interfaces, but these connections are often difficult to maintain, particularly as more systems are connected. Other systems are not linked at all and require printing data from one system and manually reentering it into another.

### **Methodological, Organizational, and Policy Issues Further Impede Collection, Analysis, and Sharing of Training Evaluation Data**

SMEs voiced concerns about the following issues:

- Aggregate measures of training quality are not comparable across schools and may be subject to misinterpretation.
- School staff might resist collecting data without a clear rationale.
- Measurement, data collection, and analysis could strain personnel resources and skills.
- The organizational culture is one in which schools generally consider performance measures to be their own concern.
- Results could be used to reduce schools’ resources.

- Existing policies on how to collect, share, and use data are insufficient.
- Policies alone are not enough to encourage data sharing; top-down guidance and incentives are needed to overcome the tendency for schools to keep data to themselves.

### **Recommendations Based on the Assessment of Army Training Data and Information Systems**

Technical impediments can be addressed in phases:

- **In the short term, the Army should help schools enhance their ability to evaluate training.** This process can be facilitated by moving to online administration of tests and surveys and by collecting different types of data within a single instrument. Manual studies should be continued until IT systems can fully support them; such studies not only provide results about the quality of training, but can yield lessons learned for designing automated collection of evaluation data.
- **In the medium term, the Army should build its ability to collect training evaluation data.** Critical to this ability will be the development of standards, which will help produce more complete, consistent, and available data. Standards are also necessary to move to service-oriented architecture (SOA). Modifications should also be made to allow training systems to collect data consistently. Web services should be modified to allow database queries.
- **In the long term, move to service-oriented architecture.** New architectures, notably SOA, have begun to emerge to support interoperability. SOA repackages the capabilities of systems into a set of “services” that can be located and used by other services, systems, or users on the network. Adopting SOA as the architecture appears to have the potential to greatly facilitate the collection and sharing of evaluation data, though not without associated risk and substantial cost. In principle, SOA would enable training systems to interoperate without prior agreement of any

kind, thereby bypassing the cost and lead time of creating specialized interfaces between systems.

**Addressing nontechnical impediments to an enterprise-wide program of evaluation is a necessary component of an implementation plan:**

- **Build end-user participation into all phases of process design.** We recommend inviting staff from proponent schools and centers as well as other organizations to participate in designing the processes used to collect and evaluate training data.
- **Establish the business case before requiring any new data collection.** Efforts should be undertaken to determine the value and feasibility of collecting various kinds of training evaluation data and to communicate the rationale for enterprise-wide evaluation efforts.
- **Develop appropriate policies to support evaluation.** The business case should be used to establish data-reporting requirements. These requirements should be crosswalked and integrated with relevant existing policies to avoid requiring overreporting. Policy should also be developed to spell out how training evaluation data will be used.
- **Provide requisite resources and incentives.** ATSC should ensure that the proponent schools and centers have the necessary resources to collect, analyze, and report evaluation data. This should include providing hardware and software for collecting, analyzing, and/or reporting data, resourcing the personnel needed to support these efforts, and providing training in analytical techniques.
- **Establish an entity to support enterprise-level training evaluation.** This office would provide analytical support to proponent schools and centers, identify relevant interoperability shortfalls and serve as a liaison to coordinate data exchange, coordinate data-collection efforts across the schools, integrate results (with input from schools) and report to ATSC and DA, and collect and disseminate lessons learned and best practices.

## Recommended Way Ahead for Army Evaluation of DL

The results of this series of studies provide information about the current state of DL and suggest that a more comprehensive program of evaluation, better supported by the Army's IT systems, could provide major benefits to TADLP. Some of the current tools are ready to be turned over to the Army for implementation. Widespread adoption of the graduate and nongraduate surveys would help individual schools and centers to systematically evaluate DL courses and provide TADLP with aggregate measures of program quality. In addition, relatively little modification would be needed to use the surveys for other forms of DL, such as blended learning or mobile training teams (MTTs).

Evaluation of learning from DL needs additional investigation. Studies conducted in one or more large-scale DL conversion courses that provide both institutional and operational training (by serving as a job aid) could yield valuable lessons and raise awareness of the value of DL to the Army training community. We anticipate that these investigations will also set the stage for measuring the impact of training on job performance.

We also recommend evaluating a broader range of training outcomes. A critical area for future study is the effect of DL on acquisition and retention of skills (in addition to retention of knowledge). An important area for future investigation is developing and testing finer-grained measures of skills that distinguish among levels of student performance. Other outcomes to assess in a comprehensive evaluation program include:

- Post-training attitudes such as self-efficacy, which can be readily added to end-of-course surveys.
- Transfer performance, i.e., behavior on the job.
- Cost and benefits of DL.

In addition to expanding the range of measures used, training evaluation should be extended to other DL approaches. For example, the Army is moving toward increased use of blended learning, as described earlier, and mobile learning (mLearning), which involves the

use of technologies such as netbooks, tablet computers, electronic book readers, personal digital assistants, and smart phones. Research efforts are needed to determine appropriate measures and methods for evaluating these forms of training.

In conclusion, improved tools and metrics for evaluating DL training can provide benefits to TADLP at multiple levels. At the student level, evaluation can enable training staff to determine student success and diagnose needs for remediation. At the course level, evaluation can show how DL affects learning and subsequent outcomes such as knowledge and skills retention and performance on the job; point to needs for improvement in course content or delivery; and determine the effect of interventions designed to enhance training quality or efficiency. At the program level, evaluation can demonstrate the value of DL and support the case for resources to meet program goals.