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 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. Permission is required from RAND to reproduce, or reuse in another form, any of our research documents for commercial use.
This product is part of the RAND Corporation occasional paper series. RAND occasional papers may include an informed perspective on a timely policy issue, a discussion of new research methodologies, essays, a paper presented at a conference, a conference summary, or a summary of work in progress. All RAND occasional papers undergo rigorous peer review to ensure that they meet high standards for research quality and objectivity.
Improving the Effectiveness of Army Distributed Learning

A Research and Policy Agenda

Susan G. Straus, Jolene Galegher, Michael G. Shanley, Joy S. Moini

As part of an ambitious plan to transform its training practices, the U.S. Army has established a large distributed learning (DL) program. This program is intended to enable the Army to provide decentralized training using one or more information and communication technologies or a combination of classroom training and technology-mediated training. The Army Distributed Learning Program, or TADLP, is a comprehensive program that is implementing DL through digital training facilities, courseware, learning management systems, and other strategies and supporting mechanisms. Under this program, the Army is in the process of implementing a plan to convert 525 training courses to DL by the year 2010.¹

DL in the Army involves (1) the delivery of individual, collective, and self-development training to soldiers and units at any time and any place; (2) through multiple means and technologies; (3) using special techniques for course design, instruction, methods of communication, and organizational and administrative arrangements (Moore and Kearsley, 1996; U.S. Department of the Army, Training Doctrine Command [TRADOC], 2003). Through DL, the Army aims to achieve five goals. These goals are to (1) train to one standard, (2) improve unit readiness, (3) reduce costs, (4) improve morale, and (5) improve training efficiency (U.S. Department of the Army, TADLP, 2003). DL offers additional benefits by enabling self-paced learning, opportunities for life-long learning, and automated tracking of student performance. At the same time, DL poses numerous challenges (e.g., Young, 1998; Welsh et al., 2003).

Creating and maintaining DL courses is costly. It may be difficult to train some kinds of skills and to establish rapport between instructors and trainees in distributed settings. Further, DL environments require that instructors develop new skills and modify their teaching strategies. Students may also need to learn to use new technologies and to take a more active approach to learning than is typical in a lecture-type classroom.

The purpose of this report is to examine the benefits and challenges of DL for learning. In particular, we review research that compares DL to classroom or residential learning (RL), which is the current standard for training. We examine the effect of DL on learning and recommend strategies for the Army to evaluate and enhance its own DL efforts.

Previous reviews of learning in DL have found that research consists largely of anecdotal reports and studies that lack comparison groups, strong experimental or quasi-experimental designs, or grounding in theoretical principles (e.g., Phipps and Merisotis, 1999; Wisher et al., 1999). At the request of Combined Armed Services Support Command of the U.S. Army, Straus et al. (2004) reviewed empirical studies published in 2000–2002 that compared DL and RL for adult learners in military and university settings. Our review focused on what students learn, or objective learning outcomes, as opposed to other aspects of effectiveness, such as student satisfaction or the costs of development or delivery.³ Our review was restricted to studies that used quantitative empirical methods, because this approach is most likely to yield reliable assessments of the relative effectiveness of DL and RL. We identified relevant studies by reviewing the contents of 27 journals and four sets of conference proceedings in education, technology, psychology, and

¹ The projected rate of courseware redesign was based on an annual target goal of 31 courses per year through FY02 and 47 courses per year thereafter (U.S. Department of the Army, TRADOC, 2001).

² The majority of studies reviewed in these reports were published between 1986–1998.

³ Many studies of DL measure satisfaction with instruction. Although satisfaction is a component of training effectiveness, it is not a measure of learning.
management as well as reports published by the Army Research Institute. To supplement our analysis, we interviewed training developers, instructors, and other training staff at five Army installations.

Key findings of this work are as follows:

• Although hundreds of studies were examined, we identified only 13 studies that measured objective learning outcomes and had unambiguous research findings.\(^4\)

• The course topics and media used for DL varied considerably, making it difficult to draw systematic conclusions about instructional media. For example, examples of media included one-way and two-way video-teletraining (VTT), Web-based instruction, computer-based training, chat, and self-study. There also were a number of different combinations of technologies used in DL (e.g., VTT alone or VTT plus chat).

• Results were consistent with other reviews that found no clear advantage for DL or RL on learning.

• Many of the studies we identified were subject to methodological limitations, including the use of convenience samples rather than random assignment, small sample sizes, low response rates or differential response rates in DL and RL conditions, and insufficient evidence about the reliability and validity of measures. These shortcomings are consistent with those reported by Phipps and Merisotis and Wisher et al. in their 1999 reviews.

These findings of no advantage for DL or RL were inconsistent with recent results reported by Sitzmann et al. (2005), who conducted a meta-analysis of studies that compared classroom instruction with Web-based instruction for adult learners. Using research published from 1996–2005, they found that Web-based instruction was superior to RL for training declarative knowledge (knowledge about facts) (77 studies) but that there was no difference between media for training procedural knowledge (knowledge about processes) (12 studies). However, like the studies reviewed by Straus et al. (2004), only a small number of studies (11 studies) in the Sitzmann et al. review were true experiments in which students were randomly assigned to training conditions. Moreover, in the experimental studies, the effects of medium were reversed, such that students performed better on tests of declarative knowledge in RL than in DL. Thus, in the quasi-experimental studies in Sitzmann et al. as well as in other reviews, results showing no differences between RL and DL or better performance in DL may have been due to factors such as selection biases. That is, these differences may be due to systematic differences in the types of students who chose RL and DL, rather than to instructional medium. The implications of these findings are particularly important for the Army to consider, given that for some courses, trainees may not have a choice about the instructional medium.

Although studies such as Sitzmann et al. (2005) are promising, internal validity in most research on learning outcomes in DL remains problematic. In other words, the experimental designs used in the preponderance of studies do not allow conclusions about whether the treatment (DL or RL) caused the observed effects on learning (see Cook and Campbell, 1979, for a discussion of threats to internal validity). This state of affairs is of increasing concern because technological capabilities are evolving rapidly, and new technologies are being deployed without an empirical foundation to inform decisions about their use. Without the knowledge that might be derived from such research, it is difficult to determine which technologies are most likely to be beneficial for particular kinds of training and how those technologies might best be employed.

Thus, we present here three sets of recommendations to help the Army meet its goals for DL. The first set of recommendations focuses on the kinds of questions that should be addressed in a well-defined program of research on the effectiveness of DL. The second set of recommendations focuses on improving the quality of research by using rigorous methods that yield credible results. Addressing the research questions we propose and using the research techniques we describe in future analyses of DL programs will strengthen the Army’s knowledge base and enable the Army to increase training effectiveness. Finally, we recommend a set of policies that are needed to support the design, implementation, and evaluation of DL in the Army.

---

\(^4\) We originally identified 33 studies that appeared to fit our criteria. After careful review, we eliminated 20 of them because they measured outcomes other than learning (e.g., satisfaction, graduation rates), consisted of the same data published in different journals or conference proceedings, or were subject to methodological weaknesses that rendered the results ambiguous. For instance, in several studies, differences in the characteristics of students in the DL and RL conditions (e.g., in pre-test scores, experience, or age) were not controlled for in the analyses of study outcomes.

---

**Establish a Research Agenda Focused on the Effects of Variations in the Design of Distributed Learning**

Many studies of DL have compared DL and RL, but we believe that focusing on DL alone, and the factors that influence its effectiveness, would be more valuable. This recommendation is based on two...
arguments. First, it is often impractical or too costly to conduct rigorous field experiments that compare results for the same course given in both RL and DL formats, which means that studies of this sort are likely to suffer from such methodological problems as lack of random assignment and small sample sizes, which undermine their value. Second, the Army is already moving to DL, in part because it may offer such advantages as the potential for increased personnel readiness and the ability to train at any time and at any place. These advantages may justify investments in DL even if performance is no better than in RL. Thus, research focused on enhancing the quality of DL—rather than contrasting RL with DL—is more likely to yield results that are of practical value to the Army.

Below, we identify specific questions for a research agenda on DL. We regard these questions as particularly important, but they should be seen as exemplars of the kinds of questions that might be asked. There are, no doubt, other specific issues that might be addressed in a comprehensive analysis of DL. Regardless of the particular issue addressed, the design of specific DL configurations, as well as corresponding research questions and measures, should be guided by theory in education and other relevant disciplines.

Use a Student-Centered Approach to Frame Research Questions

Some researchers have discussed the importance of using a student-centered approach rather than an instructor-centered approach to learning in DL. Student-centered approaches, such as the constructivist model, assume that students learn more effectively if they take responsibility for their learning and discover information on their own rather than being told what to do by an instructor or machine. Indeed, Sitzmann et al. (2005) found that in comparison with students in RL, students in Web-based courses acquired more declarative knowledge when the courses had high levels of learner control than did students in Web-based courses with low levels of learner control. In comparisons of classroom instruction with and without computer-assisted instruction, Kulik (1994) found that classes with assisted instruction produced positive effects on learning when compared with traditional instruction, especially when the computer acted as a tutor—that is, when it divided material into a sequence of small steps, sought learner responses at each step, and provided immediate feedback after each response. Because many technologies used in DL permit or require self-directed work, a model of training and the assessment of training that focuses on what students are doing as they learn seems appropriate.

Examples of research questions pertaining to self-directed learning include the following:

- What types of student-centered approaches are appropriate for different forms of DL and what are the effects on learning?
- How do variations in course structure—for example, number of assignments, frequency of deadlines, and requirements for collaboration—affect self-paced learning and students’ success in DL?
- What student characteristics influence success in self-directed learning and how should information about student individual differences affect development and delivery of DL? For instance, should the degree of course structure or learner control afforded to students be based on such characteristics as locus of control or mastery (learning) versus performance (outcome) orientations (e.g., Wang and Newlin, 2000)?

Examine Specific Technology Configurations

A wide variety of technologies, including videotapes, computer-based training, Web-based instruction, and VTT, may be used in DL. Technologies vary with respect to such characteristics as whether the instruction occurs in real time (e.g., VTT) or asynchronously (e.g., computer-based training), the depth of communication they permit between the instructor and the student (e.g., voice versus text-based media), and the degree of interactivity between the information source and the student (e.g., watching a videotape versus playing a role in a constructive or virtual simulation). Examining the effectiveness of different technology configurations for DL should provide knowledge that can be used to improve the design and outcomes of Army training programs.

Questions of this sort include the following:

- What technologies or combinations of technologies are most effective for DL, and how do these technologies interact with characteristics of the course topic, students, instructors, and aspects of the training environment?
- How can the Army make the most effective use of advanced DL technologies, such as intelligent tutoring and virtual and constructive simulations?

Examine Strategies for Facilitating Interaction in Distributed Learning

One challenge of DL for learning is that it restricts communication between instructors and students and among students. Studies of collaborative learning show that interaction among students is an important source of knowledge acquisition for a variety of learning tasks (e.g., Slavin, Hurley, and Chamberlain, 2003). A number of aspects of com-
... features of technologies for DL can create obstacles to achieving a shared understanding of course material and may disrupt opportunities for students to learn from instructors and each other.

Communicating at a distance are likely to interfere with this process (Straus and Olivera, 2000). For instance, communicating via such modes as telephone, e-mail, and online chat can require more planning or effort and are typically less elaborate or detailed than the types of interactions that occur when students and instructors are collocated. These features of technologies for DL can create obstacles to achieving a shared understanding of course material and may disrupt opportunities for students to learn from instructors and each other.

Examples of questions that address communication in DL include the following:

- What aspects of communication (e.g., frequency, media used) contribute to better outcomes?
- Under what conditions can and should collaboration between students be fostered in DL, and what mechanisms can be used to support collaboration?
- What strategies (e.g., peer tutors) can be used to overcome some of the problems of mentoring at a distance?

Examine incentives, training, and support systems for distributed learning developers and instructors

DL instruction requires different methods of designing training, delivering material, and interacting with students than does traditional classroom training. We have seen little research that delineates what skills are needed and how to train instructors to teach in DL environments. In addition, our interviews with training staff suggested that current incentive schemes are inconsistent with the requirements for the design and delivery of effective DL courses.

Examples of questions that address the development and application of these skills include the following:

- What is the effect of providing training to instructors for teaching in DL environments?
- What compensation and incentive structures are needed to motivate trainers to use appropriate instructional practices for DL?
- How do different incentive structures for organic training developers and external contractors affect efficiency in course development (e.g., through sharing and reuse of training content) and course quality?

Improve the Quality of Research on Distributed Learning

In addition to shifting the focus of its research on DL, the Army should strive to improve the quality of this research.

Use Random Assignment Whenever Possible and Measure Relevant Covariates of Learning

In studies that examine different interventions or ways of implementing DL (for instance, different technology configurations or different types of training for DL instructors), randomly assigning participants to conditions is the best way to ensure that any differences in performance between the groups can be reliably attributed to the intervention. When random assignment is impractical (e.g., in field studies of training in existing groups), it is especially important to measure and control statistically for the effects of variables that can typically affect performance in training; such variables include student cognitive ability, personality traits, experience with course material, and demographic characteristics. In the context of DL, variables such as previous experience with DL and computer self-efficacy—one’s confidence in using computers—may also affect student performance.

Use Large Sample Sizes or More Powerful Study Designs

Findings of “no difference” between DL and RL cannot be reliably interpreted to mean that DL and RL are equally effective. In the language of research, one cannot confirm the “null hypothesis.” To be confident that a finding of no difference is meaningful, one needs sufficient statistical power, which is determined, in part, by sample size. For example, if one is planning a study that compares two groups of students—one that receives RL and one that receives DL—393 students are needed in each group to conclude that an absence of difference between groups is meaningful (Cohen, 1988). Many previous studies comparing RL and DL lack sufficient power to conclude that a lack of difference between conditions is “real.”

Thus, when comparing alternative interventions, such as different technology configurations for DL, and when anticipating a small effect, it is important to use sufficiently large samples or research designs that provide adequate power. In contrast, if one expects a moderate (or larger) effect of the intervention, a much smaller sample is required. Researchers also can use study designs that require smaller samples. For instance, fewer study participants typically are needed to determine whether a single intervention

---

5 The larger the sample size, all other things being equal, the smaller the error in measurement and the greater the precision or reliability of the results (Cohen, 1988). When results are measured with greater precision, there is a higher probability of detecting the phenomenon of interest.

6 This assumes conventional standards of a small effect size (d = .2), alpha = .05, and power = .80 (Cohen, 1988).

7 For instance, if one expects a moderate effect size (d = .5) with alpha = .05 and power = .8, 64 students are needed in each group.
makes a difference (e.g., improves learning) or when the same participants experience multiple interventions than when comparing alternative interventions among different groups of participants.\(^8\)

**Use a Multivariate Approach to Measuring Training Success**

The effectiveness of training can be measured many ways. Using multiple measures can help training developers and instructors determine whether and what trainees learn from training. It is particularly important to use objective measures of learning (e.g., performance on skills tests) and to measure knowledge retention and job performance after students complete training. In particular, few studies have measured job performance following training. Other indicators of learning from training include self-efficacy (confidence in one’s ability to perform a task), commitment to training goals, and motivation (Kraiger, Ford, and Salas, 1993).

**Use Appropriate Statistical Tests to Analyze Empirical Results**

We found that in a number of published studies, incorrect statistical tests were used, thus compromising the meaning and value of results. Bonk and Wisher (2000) provide a useful guide for e-learning research intended for military training developers and planners, instructional designers, and program evaluators. They describe instructional design issues, appropriate research methods, and suggested topics for research.

**Policy Recommendations to Support Design, Implementation, and Evaluation of Distributed Learning in the Army**

To reap the benefits of an expanded research effort on the quality of DL, we recommend a number of policies for TADLP and other organizations that are responsible for training in the Army. These recommendations are based on the interviews that we conducted with training developers, instructors, and other training staff.

**Use Empirically Tested Measures of Learning and Make Results Available Within the Training Community**

Comprehensive assessment should become a routine procedure in every DL course. We suggest that TRADOC identify empirically tested measures of learning and provide a means for training developers and quality assurance personnel to access them online. These measures can then be used to create course evaluation forms that are customized to particular courses (Straus et al., 2003). In addition, making results of course evaluations available to the training community at large will provide data to support DL efforts across proponent schools.

**Develop In-House Expertise in the Development and Assessment of Distributed Learning Courses**

We suggest that the Army develop in-house expertise in the development and assessment of DL courses to avoid becoming overly dependent on contractors and to make more cost-effective use of their services. The more Army personnel understand about how to develop and evaluate DL products, the better they will be at writing appropriate contracts, conducting oversight of product development, and providing appropriate quality-assurance testing. Key practices the Army should develop to become informed consumers are described in Shanley et al. (2005, Chapter 3).

One way for the Army to develop this expertise is to use public-private partnerships in which Army training personnel work alongside contractors and eventually assume responsibility for the training function—an approach that state governments have found effective for implementing and managing information technology projects (Anderson et al., 2003).

**Enhance Opportunities for Developers and Trainers to Communicate**

Course development and evaluation are generally decentralized in the proponent schools. Our discussions with personnel at various schools suggest that the training development community would benefit from additional opportunities for formal or informal interaction. For instance, we found that DL modules developed at one school were applicable to courses being developed at other schools, but the training developers were not aware of these potential synergies. Likewise, useful information about the effectiveness of DL may not be routinely disseminated in the Army training community.

The Army already offers several mechanisms for information-sharing, including the TADLP Bulletin, the Army Training Support Center Individual Training Support Directorate DL conferences, and TRADOC Knowledge Net (TKN), which includes collaborative tools such as threaded discussions, instant messaging, and repositories where staff can post and download documents. Apparently, use of TKN has been limited, perhaps in part because of

---

\(^8\) See Cohen (1988) for recommendations regarding sample sizes needed for different study designs.
of training staff. However, we do not recommend using these outcomes as metrics because they are typically beyond the control of individual training developers or instructors. To influence the behavior of these training staff, incentives should, instead, be focused on effective processes or behaviors, such as (1) applying instructional design principles in DL course development and delivery, (2) creating SCORM-compliant courseware to facilitate reuse of content among related courses, (3) participating in communities of practice to share information about DL efforts, (4) conducting systematic evaluations of DL courses and revising courses based on the results of these evaluations, and (5) revising courses to keep both the content and the means of delivering them up to date. TRADOC should also develop metrics to assess the broader goals of DL specified above at the proponent school level.

**Summary**

Research on DL has not kept pace with the implementation of DL efforts. We recommend that the Army continue to conduct research on DL, emphasizing how to improve DL rather than comparing DL and RL. Such research might focus, for instance, on the effects of student-centered learning approaches, specific configurations of DL technologies, and strategies for facilitating interaction among instructors and students. Many other factors that could contribute to learning in DL could be examined in this way. We also encourage the Army to improve the quality of this research by conducting studies that are based on more rigorous research designs, use larger sample sizes, and examine multiple aspects of learning. Finally, we recommend providing support for effective DL development, implementation, maintenance, and improvement through organizational resources and personnel practices. Such support should include establishing mechanisms to make systematic evaluation a standard practice in DL courses. These recommendations can help the Army use DL to realize its vision of transforming training and providing lifelong learning for multiskilled soldiers.

---

*资源必须可用以更新训练，以跟上军队快速变化的运营环境。*
Bibliography


