This report is part of the RAND Corporation research report series. RAND reports present research findings and objective analysis that address the challenges facing the public and private sectors. All RAND reports undergo rigorous peer review to ensure high standards for research quality and objectivity.
Developing Army Leaders

Lessons for Teaching Critical Thinking in Distributed, Resident, and Mixed-Delivery Venues

Susan G. Straus, Michael G. Shanley, James C. Crowley, Douglas Yeung, Sarah H. Bana, Kristin J. Leuschner

Prepared for the United States Army
Approved for public release; distribution unlimited
The research described in this report was sponsored by the United States Army under Contract No. W74V8H-06-C-0001.
Executive Summary

The Command and General Staff School (CGSS) within the Command and General Staff College provides the Command and General Staff Officer Course (CGSOC) (formerly called Intermediate Level Education), which is a key component of the U.S. Army’s system for developing these and other critical knowledge, skills, and abilities in its officer corps. The Common Core, which is the first phase of CGSOC, is taught in three venues: a resident course taught at Fort Leavenworth and at satellite campuses; The Army School System (TASS), primarily for Reserve Component officers, taught by the U.S. Army Reserve Command’s 97th Brigade and its three subordinate battalions; and Advanced Distributed Learning (ADL), a web-based, self-paced course that uses interactive multimedia instruction. CGSOC consists of nine blocks of instruction taught as stand-alone modules in the resident course and organized into three phases in TASS and ADL.

In response to the interests of Army leadership, this study sought to answer the following questions about the Common Core, focusing on the 2009-2010 academic year:

- Based on current methods of evaluation, how effective is the Common Core, and to what extent are there differences among delivery venues?
- Based on current measures, how can course delivery be improved?
- How well do current methods of evaluation gauge course success and point to needed improvements?

To answer these questions, we analyzed available data from CGSS, including responses to student surveys, grades on assignments, and student characteristics. In addition, we conducted a quasi-experimental study to assess consistency in grading among faculty members.

Students in All Venues Were Generally Satisfied with the Common Core

Students in all venues generally reported that the course contributed to their learning. Students were also satisfied with other aspects of the course, including the quality of exams and other assignments and instructional delivery. While there were some statistically significant differences among venues, these differences were small and are not practically significant.

However, there were a few differences in students’ responses across venues. Students in ADL had lower ratings of the extent to which they felt that feedback enhanced their learning. In addition, some TASS and ADL students reported technical and administrative problems with the interactive multimedia instruction (IMI) portions of the course. Responses to open-ended questions revealed differences not captured in objective questions; for example, dissatisfaction of TASS students with the workload during Phase II of the course (which meets for one weekend per month for eight months) and the desire for peer interaction among ADL students.
Student Grades Were High Across all Venues; Reliability Across Graders Needs Improvement

Analysis of grades on assignments showed that average grades were generally high (ranging from 88 to 92 out of 100). There were no meaningful differences in grades across venues. Whereas grades can, in principle, reflect objective measures of learning, information about students’ levels of knowledge and skills at the outset of the course or use of a control group is needed to draw conclusions about how the course contributes to student learning. The average scores on assignments also suggest that there is grade inflation. Thus, assessments may not reflect whether students mastered the course content, and these “ceiling effects” limit the ability to discriminate among levels of performance or to observe associations with other variables.

Moreover, meaningful comparisons of performance require that faculty use the same standards and grading practices within and among venues. We conducted an exploratory, quasi-experimental study examining the consistency of grading among faculty on four assignments. Although the number of faculty who participated in the study was quite small, responses of those who did participate showed that reliability across these graders was generally quite low.

Future Evaluation Should Focus on Whether the Best Possible Outcomes Within Venues, Rather Than Equivalent Outcomes Across Venues, Are Achieved

We found few meaningful differences in students’ self-assessed learning and course grades among the three venues. However, these results should not be interpreted to mean that the venues are equally effective or that differences do not exist. Several factors make it difficult to draw firm conclusions about venue differences, including fundamental differences or confounds among venues in factors that may influence learning and attitudes (e.g., use of different learning approaches, different types of students in each venue), potential biases in surveys responses (e.g., leniency in ratings, careless responding, and survey fatigue – which likely are due to the number and length of surveys administered in the course), and low inter-rater reliability in grades.

Moreover, we would not expect the venues to have equivalent outcomes. Although the learning goals are the same, the venues are designed to support different student environments and needs. Given inherent confounds and the demanding data requirements to allow a robust comparison of venues, coupled with the expectation that the Army will have an ongoing need to provide the Common Core in different venues, we recommend that future evaluations focus on whether the best possible outcomes within venues, rather than equivalent outcomes across venues, are achieved.
Suggestions for Improvement Address Course Design and Supports As Well As Course Evaluation

Some results point to strategies to improve the Common Core. First, CGSS should continue recent efforts to examine grading practices and, if needed, provide professional development to ensure that faculty grade to a common standard. Second, CGSS should investigate the use of blended learning strategies, such as those used in their Advanced Operations Course, to provide ADL students with opportunities for instructor and peer interaction. Third, CGSS should address technical issues with access to and functions of online courseware; moving the learning management system from a dotmil to a dotcom domain, as CGSS is doing, can improve speed and reliability. The demands of Phase II on TASS students also warrant further investigation, along with consideration of mitigation strategies such as changes in course structure.

Evaluation can also be improved. Specifically, surveys used in the Common Core can be modified to obtain more diagnostic information:

- Assess self-efficacy questions before and after training in combination with individual-level explanatory variables to assess knowledge gain.
- Reduce response burden in surveys by asking fewer detailed questions and by sampling students to complete the surveys.
- Move from five- to six-point response options for close-ended questions to increase variability in responses.
- Analyze responses to open-ended questions to better understand students’ experiences in the course and to improve future surveys.
- Use web analytics, such as response time per question, to assess the quality of responses.

Finally, CGSS can conduct more comprehensive evaluations. Most important is consistently collecting information on the effect of training on job performance by conducting surveys of graduates and supervisors to evaluate whether graduates’ knowledge or performance changed after the course. Study findings also indicate the need for better infrastructure for monitoring and evaluation within CGSS to obtain consistent or reliable data through better integration among data sources, use of standardized measures within and across venues, and sufficient staff with knowledge and skills in survey development and analysis.