



INSTITUTE FOR DEFENSE ANALYSES

Phase 1 IWAR Test Results

J.D. Fletcher

January 2010
Approved for public release;
distribution is unlimited.
IDA Document D-4047
Log: H 10-000220



The Institute for Defense Analyses is a non-profit corporation that operates three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.

About this Publication

This work was conducted by the Institute for Defense Analyses (IDA) under contract DASW01-04-C-0003, Task DA-2-2896, "Technical Review and Analyses for Education Dominance," for the Defense Advanced Research Projects Agency. The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Copyright Notice

© 2010 Institute for Defense Analyses, 4850 Mark Center Drive, Alexandria, Virginia 22311-1882 • (703) 845-2000.

This material may be reproduced by or for the U.S. Government pursuant to the copyright license under the clause at DFARS 252.227-7013 (NOV 95).

INSTITUTE FOR DEFENSE ANALYSES

IDA Document D-4047

Phase 1 IWAR Test Results

J.D. Fletcher

Contents

Summary	S-1
Phase 1 IWAR Test Results	1
A. Background	1
B. Defense Advanced Research Projects Agency (DARPA) Education Dominance Program.....	1
C. Phase 1 Information Warfare (IWAR)	2
D. ACE/CID Training	3
E. IWAR-L	4
1. IWAR-L Organization	4
2. IWAR-L Support Teams	4
3. IWAR-L Testing Laboratories	5
4. IWAR-L ACE/CID Student and Fleet IT Teams	5
5. IWAR-L Schedule	6
6. Comparisons of ACE/CID Students and “A” School Students	7
7. Comparisons of ACE/CID Students and Fleet ITs	9
8. IWAR-L Results	10
F. IWAR-D	21
1. IWAR-D Organization	21
2. IWAR-D Schedule	21
3. IWAR-D Results	23
G. IWAR-A	24
1. IWAR-A Organization	24
2. IWAR-A Schedule	24
3. IWAR-A Results	25
H. Discussion	25
References	Ref-1
Glossary	GL-1

Figures

1. Example Troubleshooting Problem Summary	14
2. Example Trouble Ticket as Presented to ACE/CID Students and Fleet ITs	14
3. Ships Used During IWAR-D	22
4. Example IWAR-D Survey	22
5. Ship Used During IWAR-A	24

Tables

1. IWAR-L Daily Schedule	6
2. ACE/CID Students Compared With “A” School Students	7
3. Average Age and AFQT Scores of CID IT “A” School Students	8
4. Performance of ACE/CID and “A” School Students on CID IT Test	8
5. Comparison of Fleet ITs and ACE/CID Students	10
6. Knowledge Test Results	11
7. Knowledge Test Items on DT Content: Fleet ITs and ACE/CID Students	12
8. Percent of Troubleshooting Problems at Different Levels of Impact and Difficulty Presented on Each Day	15
9. Troubleshooting Test Results	16
10. Security Testing Results	18
11. Design and Implementation Objectives Scored 4–5 (Successful)	20
12. Design and Implementation Scores (Day 5)	20
13. Survey Responses Concerning Tasks Assigned to ACE/CID Students	23

Summary

The Defense Advanced Research Projects Agency (DARPA) Education Dominance Program is a research effort intended to substantially improve Defense training and human performance by capturing in scalable computer technology the best practices of human tutors. Among other benefits, this approach promises to significantly compress the time needed by technicians to achieve high levels of expertise. This effort focused on the Navy's Information Systems Technician (IT) rating. DARPA, through its research contractor, undertook the development of a Digital Tutor (DT) that would capture in computer technology the best instructional practices of highly capable tutors.

DT development was planned for two phases. Phase 1 ended on 30 September 2009. Initiation of Phase 2 depended on findings from Phase 1. During Phase 1, representative "A" School IT candidates were trained in cooperation with the Navy's Center for Information Dominance (CID) in a 16-week Advanced Concepts in Education (ACE) program of instruction. About 2 weeks of instruction was presented by an early available component of the (computerized) DT. The remaining 14 weeks of the course were presented by highly qualified human tutors. ACE/CID training focused on core networking competencies derived from IT "A" School content, substantial portions of five relevant "C" schools, and a careful analysis of Fleet IT requirements.

The challenge for this training was to show that its students were at least equivalent in knowledge and skill to Fleet ITs with 4–12 years experience. If this challenge was met, development of the fully computerized DT (Phase 2) was scheduled to begin 1 October 2009 and end 30 September 2011. Phase 1 Information Warfare (IWAR) testing was conducted to assess this possibility and consisted of three activities:

1. **IWAR-Laboratory (IWAR-L).** This activity compared the technical capabilities of ACE/CID students with experienced Fleet ITs in highly instrumented laboratory conditions supplemented by human observers
2. **IWAR-Dockside (IWAR-D).** This activity assessed the technical capabilities of ACE/CID students aboard docked Navy ships
3. **IWAR-Afloat (IWAR-A).** This activity assessed the technical capabilities of ACE/CID students aboard deployed Navy ships.

ACE/CID training of 12 “A” School candidates was completed on 21 May 2009. No differences of statistical or practical significance in age or Armed Forces Qualification Test (AFQT) scores were found between the ACE/CID students and typical “A” School students. CID prepared and conducted a test to ensure that ACE/CID students were receiving adequate IT “A” School instruction. Scores of the ACE/CID students on the test were significantly higher ($p < .01$) than those of a cohort of “A” School students.

A. IWAR-L

Twelve ITs were detailed from the Fleet to participate in IWAR-L testing. They averaged 8.5 years of Navy experience and 7.2 years of Navy IT experience. Overall, they had completed 41 “C” schools. Their AFQT scores were slightly higher than, but effectively equivalent to, those of the ACE/CID students.

IWAR-L was conducted in two 5-day cycles: 13–17 July 2009 and 27–31 July 2009. Six ACE/CID students and six Fleet ITs participated in Cycle 1, and the remaining six ACE/CID students and six Fleet ITs participated in Cycle 2. CID made three classrooms available for IWAR-L at the San Diego Naval Base. Two classrooms were used for testing, and one classroom housed the hardware and software needed to manage and operate systems in the testing rooms. Systems in the testing rooms were made to resemble shipboard Navy systems as closely as possible.

IWAR-L was divided into four major activities:

1. **Knowledge Testing (4 hours, Day 1).** Knowledge Testing consisted almost entirely of short-answer questions on 12 topics. The ACE/CID students received higher scores on 11 of the 12 test topics. Nine of these differences were statistically significant ($p < .05$ and $p < .01$) as were the total scores ($p < .01$) received by the ACE/CID students.
2. **Troubleshooting (13.25 hours, Days 1–3).** Almost half of IWAR-L was conducted as Troubleshooting, which was viewed as the most realistic, representative, and valid assessment of IT readiness for Fleet duty. Most IWAR-L troubleshooting problems were based on an analysis of 1,093 Casualty Report (CASREP) trouble tickets drawn from much larger Fleet Remedy databases.

The six Fleet ITs and six ACE/CID students in each cycle were divided into two three-member teams, one of which worked on virtual problems and one of which worked on physical problems. Fleet ITs and the ACE/CID students were assigned the same troubleshooting problems in the same sequence.

Overall, Fleet ITs attempted 95 troubleshooting problems and successfully solved 79 (83%) of them, with 46 (58%) of their solutions rated excellent. ACE/CID students attempted 102 troubleshooting problems and successfully solved 99 (97%) of them,

with 89 (90%) of their solutions rated excellent. The Fleet ITs made 18 harmful changes in the process of troubleshooting, and the ACE/CID students made 8 harmful changes in the process of troubleshooting. The ACE/CID students were more systematic in their troubleshooting processes as represented by 97% problems and 95% solutions verified compared to 85% problems and 77% solutions for the Fleet ITs.

3. **Security Testing (7 hours, Day 4).**¹ For this exercise, the physical systems in each testing room were pre-loaded with about 100 typical shipboard security vulnerabilities. Fleet and ACE/CID teams were to restore the systems to Fleet security specifications while avoiding changes that did not conform to Fleet regulations. No team did well on this portion of IWAR-L. Fleet teams scored 149 (35%) of the total points possible, and ACE/CID teams scored 95 (23%) of the total points possible. These differences do not differ significantly ($p < .05$).
4. **Design and Implementation (7 hours, Day 5).**² For this exercise, Fleet ITs and ACE/CID students were to assemble a system that met 24 objectives. Overall, Fleet ITs achieved 15 of these objectives, and ACE/CID students achieved 14 of these objectives. Neither group did well in scoring. Fleet ITs scored 113.5 (47%) points out of 240 possible, and ACE/CID students scored 84.5 (35%) points out of 240 possible. These differences were not statistically significant ($p < .05$).

B. IWAR-D

IWAR-D consisted of two 5-day cycles, 3–7 August and 10–14 August 2009. Two DDGs were made available for both cycles: the *USS Preble* (DDG 88) and the *USS Milius* (DDG 69). Three ACE/CID students were assigned to each ship in each cycle. Data consisted of student task diaries and surveys completed by supervisors.

Overall the ACE/CID students completed about 38 tasks during IWAR-D. The number of tasks was affected by shipboard schedules and activities. Most tasks assigned to the ACE/CID students were either Trouble Ticket responses or Emerging Problem tasks, which were those that were found while correcting other problems. The ACE/CID students solved about 87% of tasks assigned to them successfully. Some tasks were not solved because the ACE/CID students were not permitted access to a component or because the time allocated for the day's IWAR-D testing ran out.

Most of the contributions of the ACE/CID students were rated as substantial or essential. Two instances occurred in which the ACE/CID students solved problems that shipboard ITs had

¹ In the morning, half hour of orientation followed by 3.5 hours of the exercise in the morning for two 6-member Fleet teams working in the separate laboratory rooms. In the afternoon, a half hour of orientation followed by 3.5 hours of the exercise for two 6-member ACE/CID teams working in the separate laboratory rooms.

² 3.5 hours morning testing and 3.5 hours afternoon testing.

not been able to correct either with their own solutions or with solutions they had received from the Space and Naval Warfare Systems Center (SPAWAR), San Diego.

C. IWAR-A

IWAR-A was limited to observations of three ACE/CID students who served aboard the *USS Vandegrift* (FFG-48), 17–20 August 2009. IWAR-A data consisted of student task diaries and a narrative report from the Senior Chief who served as their monitor.

The ACE/CID students were assigned system-wide tasks that are typically addressed by senior ITs. These tasks were essential for successful operation of the ship's IT systems and would have created debilitating results if done incorrectly. All were reported to have been accomplished correctly. All three ACE/CID students participated equally in technical activities.

D. Discussion

The ACE/CID training challenge appears to have been met. On this basis, Phase 2 of the DARPA program was initiated. If successful it should advance the state of training art and practice and contribute significantly to Navy operational capabilities.

Phase 1 IWAR Test Results

A. Background

The benefits of tutorial—one tutor for (ideally) one student—instruction have long been known, understood, and documented (e.g., Bloom, 1984; Graesser & Person, 1994; Graesser, D’Mello, & Cade, in press). Tutorial instruction provides a variety of demonstrated benefits (e.g., interactivity, individualization, and engagement) and has also been suggested as a way to reliably compress the many years needed to develop technical expertise into substantially shorter periods of time (Fletcher, 2001, 2009). However, with the exception of a few critical skills, it has been viewed as an economic impossibility (Scriven, 1975).

Affordable computer technology and adaptive computer-based instruction promise, through their capabilities for massive scalability, to make tutorial instruction affordable (Fletcher & Chatham, 2010). Research and development (R&D) that attempts to realize this promise has been supported since the late 1960s, much of it sponsored by the U.S. Navy (Coulson, 1962; Fletcher & Rockway, 1986; Ford, Slough, & Hurlock, 1972).

B. Defense Advanced Research Projects Agency (DARPA) Education Dominance Program

Most research on computer-based tutoring has attempted to apply theories and approaches developed by cognitive and computer scientists (e.g., Sleeman & Brown, 1982; Psotka, Massey, & Mutter, 1988; Luckin, Koedinger, & Greer, 2007). The DARPA Education Dominance Program has taken another, more pragmatic approach. Noting that large differences can be observed and measured in the learning results achieved by different teachers, DARPA, through its research contractor, has chosen to examine and capture in computer technology the best practices of expert human tutors. The aim is to create a Digital Tutor (DT) that achieves or even exceeds the “2-Sigma” challenge laid down by Bloom (1984)—a challenge to produce two standard deviations of difference in learning over more conventional approaches. DARPA is focusing this effort on training that is intended to prepare sailors for the Information Systems Technician (IT) rating. Roughly, a two standard deviation difference would raise the performance of IT technicians at the 50th percentile to that of the 98th percentile.

DT development was planned for two phases. Phase 1 ended 30 September 2009. During Phase 1, representative “A” school IT candidates were to be trained using DT content and

methodology in a 16-week Advanced Concepts in Education (ACE) program. This training was completed in May 2009.

C. Phase 1 Information Warfare (IWAR)

IWAR testing was to determine if Phase 1 training had produced IT technicians who have the technical knowledge and skills of technicians who have 4–12 years experience in the Fleet. If IWAR showed that the DARPA approach had met this objective, Phase 2 would begin development and testing of the fully computerized DT on 1 October 2009. Similar IWAR testing would be conducted at the conclusion of Phase 2.

As indicated by the data presented in this report, Phase 1 successfully met and, in fact, exceeded its objectives. Phase 2 of the DARPA program was initiated on 1 October 2009.

The training and testing required for this effort were made possible through cooperation with the Center for Information Dominance (CID) Corry Station, Pensacola, Florida. CID supplied (1) students from the “A” school pool at Corry Station, (2) space and facilities to train them with DT content and techniques at the CID Detachment, Monterey, California, and (3) space and facilities for IWAR assessment at the CID detachment at the San Diego Naval Base. The Navy IT candidates who participated in this research are designated as ACE/CID students in this report.

Three assessment activities were undertaken for Phase 1 IWAR testing:

1. **IWAR-Laboratory (IWAR-L).** Knowledge and performance assessment of the ACE/CID students compared with experienced Fleet ITs. Knowledge assessment was accomplished using written, short-answer response paper and pencil tests. Performance testing was conducted with both groups operating, troubleshooting, securing, designing, and assembling Fleet systems and equipment in highly instrumented laboratory conditions supplemented by human observers.
2. **IWAR-Dockside (IWAR-D).** Dockside performance of the ACE/CID students with shipboard systems and human observers, as permitted by shipboard conditions.
3. **IWAR-At Sea (IWAR-A).** At-sea performance of the ACE/CID students with shipboard systems and human observers, as permitted by deployment and shipboard conditions.

This report describes these three activities and summarizes the findings. The testing provided a controlled and complete comparison between ACE/CID students and Fleet ITs. IWAR-L and IWAR-D testing took place in San Diego from July–September 2009. IWAR-A took place at sea during September 2009.

D. ACE/CID Training

ACE/CID training began in January 2009. Most of the training was presented by about two dozen human tutors who possessed expertise in one or more IT subdomains covered by the training. Half of these tutors were selected through extensive interviews that sought to ascertain instructional expertise (demonstrated in half-hour tutoring “auditions”) and subject-matter mastery of an IT subdomain needed for the ACE/CID training. Another quarter of the instructors had written a definitive, advanced text for an appropriate IT subdomain and were then selected for their tutorial expertise. The final quarter of instructors were selected on the recommendation of qualified organizations that identified them for their tutorial and IT subdomain expertise.

Per plan, an early available portion of the (computerized) DT presented 1–2 weeks of the ACE/CID training. Time of use depended on that needed by individual students to master the material. The goal was for the ACE/CID students to receive 1:1 tutoring from the human tutors for the remaining instruction. Major exceptions were a 1-week hardware laboratory, which was taught 1:4; a half week of networking, which was taught 1:8 because of limited tutor availability; and 9 students who received 1:2 tutoring for about half of their instruction, again due to limited availability of the (human) tutor.

About 30% of ACE/CID instructional content was based on a careful analysis of Fleet IT requirements undertaken in 2008. The remaining two-thirds of the instructional content was derived from topics covered by the 11-week CID IT “A” school and the core material covered in the following 5 “C” schools:

1. Journeyman-Networking Core (JNETCORE)
2. Advanced Network Analyst (ANA)
3. Information Systems Security Manager (ISSM)
4. Network Security Vulnerability Technician (NSVT)
5. Navy Tactical Command Support System (NTCSS) Manager.

ACE/CID objectives focused on developing core competencies in networking (i.e., workstation and server software, routing and switching, operating systems, and so forth). This training was also designed to provide students the knowledge, understanding, and general problem-solving abilities that would enable them to learn and adapt to new systems and configurations quickly.

ACE/CID training covered the following basic topics:

- Hardware: Windows Basics
- Windows 2003 Server: Primary Services
- Windows 2003 Server: Applications
- Fleet Networking, Routing, and Switching
- Switching (CISCO/Alcatel)
- Radio Frequency (RF) Communications
- Network Security
- Programming.

ACE/CID training took place from 26 January to 21 May 2009. Three days in week 17 were added to make up for holidays during the 16-week period. Twelve students completed the ACE/CID training. No students were dropped from the program because of academic reasons.

E. IWAR-L

1. IWAR-L Organization

IWAR-L provided the first and most extensive testing for Phase 1 of the DARPA program. It compared the knowledge and skills of the 12 ACE/CID students who had completed ACE/CID training with the knowledge and skills of 12 Fleet ITs who, as it turned out, had 3–12 years of IT experience. The Fleet ITs were detailed from large-deck and small-deck ships.

IWAR-L was undertaken and completed in two 5-day cycles: 13–17 July and 27–31 July 2009. Six ACE/CID students and six Fleet ITs participated in Cycle 1. The remaining six ACE/CID students and six Fleet ITs participated in Cycle 2.

2. IWAR-L Support Teams

Three teams supported IWAR-L testing:

- **White Team.** The White Team conducted the IWAR-L exercises. White Team members organized participants for IWAR trials, ensured that they observed trial parameters and protocols, managed information communication, and participated in recording and scoring performance data. IWAR-L was especially fortunate in obtaining the services of the White Team leader, two CID instructors, and three Fleet Systems Engineering Team (FSET) members for this work.
- **Red Team.** The Red Team consisted of a highly experienced Senior Chief, who was assisted by equally capable FSETs. The Red Team designed and set up systems for

IWAR-L security testing and helped assess the performance of participating ITs in correcting security vulnerabilities.

- **Technical Team.** The Technical Team was provided by the research contractor and was responsible for the proper management and operation of the IWAR-L hardware and software. Commendably, especially for such a complex integration of many different systems, IWAR-L testing did not experience any major technical disruptions during its two 5-day cycles of exercises.

3. IWAR-L Testing Laboratories

IWAR-L participants worked in two large separate laboratory rooms (each approximately 20 ft. × 30 ft.) provided by CID on the San Diego Naval Base. Each laboratory room contained three Information Technology systems: two identical virtual systems that participating ITs operated via keyboard and monitor displays and one physical system with cables and other hardware that participants operated manually. The laboratory systems were designed to mirror typical shipboard Information Technology equipment (servers, routers, switches, and so forth). A third classroom housed the hardware and software that the Technical Team needed to manage and operate systems in the testing rooms.

The two virtual systems consisted of real server and networking software (e.g., COMPOSE 3.0 running on three servers with a backbone) but without the capacity of a full hardware implementation. Three workstations were used to operate each of the virtual systems. The physical laboratory system had a server rack with four servers, one UNIX system, two backbone switches, and four edge switches. Connected to the rack were three workstations on a network made to resemble shipboard systems as closely as possible.

The Technical Team was responsible for injecting troubleshooting problems and correcting any system technical issues that arose. The Technical Team ensured that the systems could be restarted quickly—when and if that was required. Spare hardware for every system component was available, along with spare disks with cloned images for the servers so that problems with the server rack would only take a few minutes to resolve.

The laboratories were also instrumented with video cameras and microphones. All activity on the monitors was time stamped for later review and analyses using tools already developed for ACE/CID training.

4. IWAR-L ACE/CID Student and Fleet IT Teams

For much of IWAR-L testing, the six participants in the two laboratory rooms were divided into two three-member teams. Simultaneously, one team worked on virtual systems to solve system problems, and the other team worked on physical problems. Members of these

teams were assigned by the White Team. For the Fleet teams, one member was an IT1 with 11+ years of experience, a second member was an IT2 with 6+ years of experience, and the third member was an IT2 with less than 6 years' experience. For the ACE/CID teams, one member was assigned from the top third of the ACE/CID students, one member from the middle third, and one member from the lower third.

5. IWAR-L Schedule

Table 1 shows the daily schedule followed for each of the two 5-day IWAR-L cycles.

Table 1. IWAR-L Daily Schedule

Day 1	Day 2	Day 3	Day 4	Day 5
<ul style="list-style-type: none"> • Introduction to IWAR testing (0.5 hour) • Knowledge Test Part 1 (2 hours) Break (30 min) <ul style="list-style-type: none"> • Knowledge Test Part 2 (2 hours) Lunch <ul style="list-style-type: none"> • Familiarization with laboratory systems (1 hour) • Troubleshooting very easy to medium difficulty problems (2 hours) 	<ul style="list-style-type: none"> • Troubleshooting easy, medium, and difficult problems (6 hours) 	<ul style="list-style-type: none"> • Troubleshooting medium to very difficult problems (5.25 hours) 	<ul style="list-style-type: none"> • Security: Identification and correction of security vulnerabilities • Morning orientation (0.5 hour) • Morning Fleet team testing (3.5 hours) Lunch <ul style="list-style-type: none"> • Afternoon orientation (0.5 hour) • Afternoon ACE/CID team testing (3.5 hours) 	<ul style="list-style-type: none"> • Design and implementation: Given objectives, design and implement an IT system • Morning testing (3.5 hours) Lunch <ul style="list-style-type: none"> • Afternoon testing (3.5 hours)

As Table 1 shows, the IWAR-L schedule involved the following activities:

- Day 1
 - Introduction to IWAR testing (0.5 hour)
 - Two-part written test of IT knowledge (4 hours)
 - Familiarization with the laboratory systems (1 hour)
 - Troubleshooting very easy to medium difficulty problems (2 hours)
- Day 2
 - Troubleshooting easy, medium, and difficult problems (6 hours)
- Day 3
 - Troubleshooting medium, difficult, and very difficult problems (5.25 hours)

- Day 4
 - Finding and correcting security vulnerabilities. In the morning, a half hour of orientation was followed by 3.5 hours of the exercise in the morning for two 6-member Fleet teams working in the separate laboratory rooms. In the afternoon, a half hour of orientation was followed by 3.5 hours of the exercise for two 6-member ACE/CID teams working in the separate laboratory rooms.
- Day 5
 - System design and implementation (3.5 hours in the morning and 3.5 hours in the afternoon). Six-member Fleet and ACE/CID teams working in the separate laboratory rooms.

To ensure standardization of treatment, a member of the White Team introduced each day’s session with a written “script” that was read to the participating ITs. The same troubleshooting problems (Days 1–3) were given to both groups in exactly the same sequence. Different troubleshooting items were presented in the two cycles, but they followed the same sequence of difficulty and impact in both cycles.

6. Comparisons of ACE/CID Students and “A” School Students

Thirty-two students (22 males and 10 females) who had been selected for IT “A” school training were offered as candidates for ACE/CID training. Fifteen were selected (11 males and 4 females). The candidate who had the highest Armed Forces Qualification Test (AFQT) score was not selected for ACE/CID training, and a student who required a waiver for admission to “A” school was selected. Two of the students chosen for ACE/CID training had learned English as a second language. None of the students selected reported any IT experience before their Navy enlistment.

Table 2 provides the data available on the 15 students selected for ACE/CID training and on the remaining “A” school students who were not selected.

Table 2. ACE/CID Students Compared With “A” School Students

	Age		AFQT Scores	
	ACE/CID Students	Remaining “A” School Students	ACE/CID Students	Remaining “A” School Students
Average	20.8	19.6	71.3	69.8
Std dev	2.7	2.2	11.8	11.5
N	15	17	15	17

The data on age and AFQT scores for students selected for ACE/CID training did not differ statistically from those who were not selected. Further, as shown in Table 3, the ACE/CID students shown in Table 2 did not differ statistically nor did they appear to differ practically from the 2008 and 2009 populations of “A” School students.

Table 3. Average Age and AFQT Scores of CID IT “A” School Students

	Age of CID IT “A” School Students		AFQT Scores of CID IT “A” School Students	
	FY2008	FY2009	FY2008	FY2009
Average	20.9	21.1	72.8	73.3
Std dev	2.8	3.2	12.7	12.4
N	517	1,172	517	1,155 ^{See Note}

Note for Table 3: Scores not available for all 2009 students.

As ACE/CID training progressed, CID command wanted to be assured that its students were receiving adequate levels of IT instruction. Accordingly, CID developed a test containing essay, diagram, and multiple-choice questions that were administered to ACE/CID students and “A” school students. The test covered the 5 weeks of network instruction received by the “A” school students and the 2 weeks of similar material received by the ACE/CID students. Table 4 shows the results of this testing.

Table 4. Performance of ACE/CID and “A” School Students on CID IT Test

	Essay		Diagram		Multiple-Choice		Totals	
	ACE/CID	“A” School	ACE/CID	“A” School	ACE/CID	“A” School	ACE/CID	“A” School
Mean	49.7	21.5	13.3	6.575	14.7	11.6	77.7	39.7
Std Dev	8.9	12.65	3.2	3.77	1.6	3.5	11.8	18.7
N	15	20	15	20	15	20	15	20
Probability	p < .01		p < .01		p < .01		p < .01	
Effect Size	2.5		1.9		1.1		2.36	

All 15 ACE/CID students took the CID IT test (3 students were later dropped from the ACE/CID training program for non-academic reasons). The differences on all the subtests and the total test scores are statistically significant ($p < .01$) in favor of the ACE/CID students. These results suggest that the ACE/CID students had successfully learned standard “A” school networking content during their first 2 weeks of training and before moving on to content drawn from “C” schools and Fleet requirements. Their average score was almost twice that of the “A” school students.

Effect sizes are reported in Table 4 and elsewhere in this report. Increasingly, effect sizes are being required and reported for single research studies. These statistics, as their name suggests, are used to go beyond the point comparisons of statistical significance and express the magnitude of treatment or experimental effect along a standardized continuum.

An effect size is usually calculated as the difference between two means divided by their pooled standard deviation. It therefore expresses treatment effect magnitude as a standard deviation. An effect size of 0.20 standard deviations is considered small. Roughly, it may be viewed as improving performance from the 50th to about the 58th percentile of performance. An effect size of 0.50 is generally considered to be medium and can be viewed as an improvement from the 50th to about the 69th percentile of performance. An effect size of 0.80 is considered large and can be viewed as an improvement to about the 79th percentile of performance (Cohen, 1988).

One challenge for the DARPA Education Dominance Program was expressed in these units. Based on Bloom's (1984) often cited challenge, it assumed a goal of making a two standard deviation improvement over current "A" school IT education and training. This goal is ambitious, although it is not a ceiling. Research may aspire to whatever effect size seems feasible.

Notably, the effect size of 2.36 on the CID test in favor of the ACE/CID students (see Table 4) well exceeds the two standard deviation goal. Given the results reported from the Knowledge Test (see Table 6), which compared performance of ACE/CID students with Fleet ITs, it seems likely that if "A" School students had taken the test instead, the effect size for ACE/CID training would have exceeded Bloom's challenge.

7. Comparisons of ACE/CID Students and Fleet ITs

Table 5 shows data comparing the background of ACE/CID students with that of the Fleet ITs who participated in IWAR-L. All Fleet ITs were rated IT2 or IT1. One student had only 3 years of IT experience but had graduated at the top of his "A" school class. Median years of Fleet IT experience was 8. The Knowledge Test identified one of the Fleet ITs who had spent all his time in communications and had virtually no experience in network technology. He was replaced by another (from the same ship) who had 7.5 years of Fleet network technology experience.

Given these data, it seems reasonable to conclude that the Fleet ITs detailed for IWAR-L were representative.

Table 5. Comparison of Fleet ITs and ACE/CID Students

Comparisons	Fleet ITs	ACE/CID Students
Average Age	27.5	20.8
Average Years Navy Experience	8.5 (range 3–18)	< 1
Average Years IT Experience	7.2 (range 3–12)	< 1
Average AFQT Scores	69.9	69.0
Total “C” Schools Completed	41	0 ^a
Total Navy Certifications	9	0
Male/Female	9M/3F	8M/4F

Note for Table 5: Some “C” school material was included in ACE/CID training.

8. IWAR-L Results

IWAR-L was divided into the four major activities described earlier:

1. Knowledge testing (4 hours – Day 1)
2. Troubleshooting (13.25 hours – Days 1–3)
3. Security testing (3.5 hours – Day 4)
4. Design and implementation (7 hours – Day 5).

IWAR-L results are reported in accord with these four activities.

a. Knowledge Testing (Day 1)

The written Knowledge Test consisted almost entirely of short-answer questions on the following 12 topics:

1. Hardware/Windows Basics
2. Client Support Fundamentals
3. Windows Server Fundamentals
4. Windows Domain Name System (DNS) Servers
5. Active Directory
6. Exchange
7. Group Policy
8. CISCO Internetwork Operating System (IOS)
9. Open Shortest Path First (OSPF) Protocol
10. Switching
11. UNIX Operating System
12. Network Security.

Knowledge Test items were “vetted” for relevance, correctness, and fairness by four members of the Institute for Defense Analyses (IDA) technical staff, two CID instructors, and an FSET.

Knowledge Test Procedures

After a brief introduction to the IWAR exercises and the Knowledge Test, both groups of 12 Fleet ITs and 12 ACE/CID students took the test together during the morning of Day 1. The test was administered as a “closed book” test in 2 parts of 2 hours each, separated by a half hour break. A White Team member was present at all times to answer questions and to act as a test proctor.

Knowledge Test Results

Table 6 shows the performance of the two groups on each Knowledge Test topic and then overall. The questions were scored 0–2, with 0 for an incorrect response and 2 for a correct response. Partial credit of 1 was awarded in accord with an anchored scoring rubric. As can be seen from the scores, the Knowledge Test was challenging for all participants.

Table 6. Knowledge Test Results

Topic	No. of Items	Fleet ITs		ACE/CID Students		Probability	Effect Size
		AVG	Std Dev	AVG	Std Dev		
Hardware/Windows Basics	10	11.7	5.1	14.7	2.7	p < .05	0.74
Client Support Fundamentals	15	12.5	7.4	21.8	4.3	p < .01	1.54
Windows Server Fundamentals	17	11.8	5.4	17.9	3.8	p < .01	1.30
Windows DNS Servers	12	3.6	3.7	11.9	3.5	p < .01	2.31
Active Directory	10	2.0	3.1	6.3	3.0	p < .01	1.42
Exchange	10	8.7	3.0	10.9	4.1	ns	0.63
Group Policy	12	3.8	2.2	6.1	1.7	p < .05	0.87
CISCO IOS	11	4.4	4.9	15.4	5.1	p < .01	2.11
OSPF Protocol	11	5.4	4.8	13.9	2.3	p < .01	2.16
Switching	12	8.9	6.1	16.4	3.8	p < .01	1.27
UNIX Operating System	9	4.9	3.3	1.6	2.1	p < .01	-1.12
Network Security	10	9.0	3.1	9.8	3.3	ns	0.26
Total Scores	139	86.7	43.9	146.7	68.0	p < .01	1.02

Note for Table 6: ns = not significant at p < .05.

As shown in Table 6, the ACE/CID students scored significantly higher on 9 of the 12 test topics and significantly higher overall, with a total score of 146.7 compared to the Fleet

total score of 86.7 and an effect size of 1.02. Notably, however, their scores for Security were about the same as those for Fleet ITs, and the Fleet ITs outscored the ACE/CID students in the UNIX subtest. About half of the ACE/CID students missed substantial portions of the UNIX training, and absence of that training shows in their scores. It was the only topic for which any ACE/CID students received a score of 0, as 6 of them did.

Scores for individual Fleet ITs were spotty. Individual ITs did well in some topics but answered almost no questions in others. The ACE/CID students were more uniform in their scores, with very few 0 for any topics, in contrast to individual Fleet ITs who either seemed to know a topic well or not at all.

If the Knowledge Test was used to determine whether the ACE/CID students were at least equal in IT knowledge to experienced Fleet ITs, it seems reasonable to conclude, conservatively, that they are. Overall, they outscored the Fleet ITs by about 68%, for an effect size of 1.02.

Whether the same content and methodology can be implemented successfully in the DT remains to be determined in Phase 2 of the DARPA program. However, a portion of the DT was available and used for ACE/CID training. The IT content presented by the DT can be matched to 21 items on the Knowledge Test (again scored 0–2). As shown in Table 7, ACE/CID students produced Knowledge Test scores for these 21 items that were significant both statistically and practically, with an effect size of 1.73. These results suggest that the digitization of ACE/CID training is feasible and that results from Phase 2 IWAR may be similar to those observed in Phase 1.

Table 7. Knowledge Test Items on DT Content: Fleet ITs and ACE/CID Students

Number of Items	Fleet ITs		ACE/CID Students		Probability	Effect Size
	AVG	Std Dev	AVG	Std Dev		
21	18.0	8.2	28.75	3.2	$p < .01$	1.73

Note for Table 7: 12 ITs and 12 ACE/CID students.

b. Troubleshooting (Days 1–3)

Troubleshooting exercises took up about half of IWAR-L time. These exercises were expected to provide the most realistic, valid, and representative assessment of the IT capabilities needed in the Fleet. The laboratory was the most appropriate setting for these exercises because it allowed the introduction of catastrophic and intractable problems, such as losing an Exchange database, cross-connected switches, looped hubs, viruses, and corrupted Active Directory. These

problems occur in the Fleet, but their introduction for experimental testing would be neither practical nor permitted in a shipboard environment.

Most IWAR-L troubleshooting items were drawn from an analysis of 1,093 Casualty Report (CASREP) trouble tickets in Fleet Remedy databases. The problems were carefully developed and tested by the technical team; vetted for correctness, fairness, representativeness, and relevance by IDA, CID, and FSET technical specialists; categorized for problem difficulty and impact on ship operations; summarized in a standard format; and presented to IWAR-L participants in the form of trouble tickets.

Difficulty of troubleshooting items was rated on a 5-point scale ranging from very easy (routine problems that could be solved by the average “Power User”) to very difficult (problems that could only be solved by the most proficient and experienced IT professionals). Impact was rated on a 10-point scale keyed to the number of IT systems (scope) affected by the problem and the impact of the problem on ship’s operations (severity). Impact ranged from minimal (one system affected, one specific operation impeded but not halted) to catastrophic (all systems affected, overall functionality severely compromised). Time was also included to resolve the problems. It was based on subject matter expert (SME) estimates of the time needed to solve a problem of given difficulty by an average IT capable of solving it (e.g., the time needed by a very proficient IT to solve a very difficult problem).

Figure 1 shows an example Troubleshooting Summary used by the Technical Team to organize and, using the injection script, set up problems. Trouble tickets were given to IWAR participants as statements of the problem to be solved. Figure 2 shows the trouble ticket associated with the problem summarized in Figure 1. ITs were instructed to complete the trouble ticket form and include a description of the problem solution and the steps taken to solve it.

Troubleshooting System Familiarization

Each group of ITs was given an hour to become familiar with the laboratory systems. Instructions were read from a script by members of the White Team, who were assisted by members of the Technical Team in conducting the familiarization. Familiarization, along with all other IWAR-L activities, was recorded by video.

Troubleshooting Procedures

Troubleshooting sessions were initiated with a problem expressed as a Trouble Ticket given to each of the two ACE/CID and Fleet teams. The same sequence of troubleshooting problems was presented to the ACE/CID students and to the Fleet ITs. When the team members finished solving a problem to their satisfaction, or they gave up, they requested a new one.

Scenario:	TS-SV-GC-30		
Concept Tested:	IP configuration and Troubleshooting		
General Description			
Add a static route for the 172.16.0.0/30 network to point to 172.16.1.254			
Injection Script			
<ol style="list-style-type: none"> 1. Log on to EX01 as the proctor admin account (proctor) 2. Open a command prompt 3. In the command prompt, enter the following command: route add 172.16.0.0 mask 255.255.252.0 172.16.1.254 -p 4. To test, try to ping WKS01. If all is configured correctly, this will fail. 			
Problem Symptoms			
<ul style="list-style-type: none"> • LT Sulu complains he is not receiving email. 			
Preferred Solution(s)			
<ul style="list-style-type: none"> • Delete the static route on EX01 			
Impact:	If clients cannot connect to EX01, they will not be able to send or receive email.		
Impact Rating:	7-High	Difficulty:	Very Hard
Time to Resolve:	30 minutes		

Figure 1. Example Troubleshooting Problem Summary

Trouble Ticket

Day 3, July 29, Time (Start/End) _____ Team: _____

Problem Symptom: Lt Sulu complains he is not receiving email

Problem Solution:

Key Solution Steps:

Figure 2. Example Trouble Ticket as Presented to ACE/CID Students and Fleet ITs

The troubleshooting problems were drawn from 150 scenarios prepared for IWAR-L. As Figure 1 shows, they were categorized by impact, difficulty, and time to solve. Although different scenarios were presented in the two IWAR-L cycles, sequencing and scenario selection were kept the same for the proportion of concepts tested, the levels of difficulty, and the impact for the two cycles. Selection and sequencing of scenarios was done by IDA and was known only to IDA before the scenarios were presented.

Outside access was carefully controlled for the troubleshooting sessions. Cell phones, email, and access to Google were not allowed. However, an attempt was made to mirror ship-board conditions as closely as possible. Reference materials for the systems were available in each room, and the participating ITs were permitted to bring and consult their own notes, references, and troubleshooting guides.

Table 8 shows the percentages of troubleshooting items at each impact and difficulty rating sequenced for Days 1–3.

Table 8. Percent of Troubleshooting Problems at Different Levels of Impact and Difficulty Presented on Each Day

Days	Impact Rating Percent					Difficulty Rating Percent				
	1–2	3–4	5–6	7–8	9–10	Very Easy	Easy	Medium	Hard	Very Hard
1	25	25	25	25	–	30	30	40	–	–
2	20	20	20	20	20	10	20	40	30	–
3	10	25	25	20	20	–	20	30	30	20

Troubleshooting Results

Table 9 shows the results from the troubleshooting exercises for the four ACE/CID teams, the four Fleet teams, and for all ACE/CID and Fleet teams combined. These data suggest, again, that the ACE/CID students did well. They solved about 25% more problems than the Fleet ITs solved.

Each problem attempted by each of the four teams (two ACE/CID and two Fleet teams) was scored by three observers: two members of the White Team, who observed troubleshooting directly as proctors, and one member of the Technical Team, who observed troubleshooting indirectly via a bank of multiple-image monitors, one of which tracked and displayed all keyboard inputs from the participating ITs. At the end of each day for each observed problem, the three observers discussed all the ratings they had assigned and the reasons for assigning these ratings.

Table 9. Troubleshooting Test Results

Cycle	Team	No. Correctly Solved	No. of Correct Rated 5 (Excellent)	No. Harmful Changes (3 to 5)	Percent Problems Verified	Percent Solutions Tested
1	Fleet (1-1)	19	12	4	79%	53%
	Fleet (1-2)	27	22	4	81%	81%
	ACE/CID (1-1)	29	27	2	100%	97%
	ACE/CID (1-2)	25	20	2	88%	92%
2	Fleet (2-1)	20	12	4	90%	85%
	Fleet (2-2)	13	11	6	91%	92%
	ACE/CID (2-1)	23	21	2	96%	91%
	ACE/CID (2-2)	22	21	2	100%	100%
All Fleet (95 attempted)		79 83%	46	18	85%	77%
All ACE/CID (102 attempted)		99 97%	89	8	97%	95%
All Fleet vs. All ACE/CID significance		p < 0.01	p < 0.01	p < 0.01	p < 0.01	p < 0.01

Note for Table 9: *The numbers in the “Team” column headings indicate the team and the cycle (i.e., Fleet (1-1) = Fleet IT Team 1/Cycle 1; ACE/CID (2-2) = ACE/CID Team 2/Cycle2).*

Problem solutions fell into three categories: correct, incorrect, or partially correct. Correct solutions received a score of 5, and incorrect solutions received a score of 0. Problems in between were given a score of 1–4, depending on the rater’s judgment.

The three observers assigned exactly the same scores to about 75% of the problems. If, after discussion, the observers could not reach agreement on a rating, the reported score was the average of the three separate ratings. Among all ratings for all troubleshooting problems, only four differed by more than one point on any scale.

Changes made in the course of problem solving that would be harmful to IT operations and not returned to their original states were also noted and rated on a 5-point scale, with 5 as most harmful. Many trouble tickets found in Fleet databases resulted from such changes made deep in the systems—changes made when troubleshooting other problems and then not restored. ACE/CID training had emphasized the importance of following proper procedures in troubleshooting (i.e., understanding before acting). This emphasis is evident in the “Harmful Changes”

ratings contrast between Fleet ITs, who had a total of 18 harmful changes, and ACE/CID students, who had a total of 8 harmful changes.

The emphasis on process is also evident in the data on the percent of problems verified and the percent of solutions tested. ACE/CID students were taught to verify trouble ticket problems and symptoms carefully before beginning to solve them. They were also taught to test solutions with equal care to ensure that they had genuinely solved the problem before concluding that they had done so. Again, this emphasis on problem-solving procedure paid off in the ACE/CID scores and the percent of problems (97%) and solutions (95%) they verified vs. the percent of problems (85%) and solutions (77%) the Fleet ITs verified.

Finally, the Fleet ITs attempted 95 total troubleshooting problems, of which 79 (83%) were solved correctly. The ACE/CID students attempted 102 troubleshooting problems, of which 99 (97%) were solved correctly. The systematic process for problem solving used by the ACE/CID students seems evident from data showing that almost twice as many of their solutions (89) were rated excellent as compared to those of the Fleet ITs (46).

All the differences between overall Fleet and overall ACE/CID performance were significant ($p < 0.01$).

c. Security (Day 4)

Realistic Security testing required the use of the single physical system available in each laboratory room. For that reason, only two teams could participate in the Security exercise at one time. The day was organized so that the two Fleet teams worked on the Security problems in the morning and the two ACE/CID teams worked on the problems in the afternoon—4 hours for each group.

Security Testing Procedure

The Red Team preloaded the physical systems in each room with about 100 typical ship-board security vulnerabilities (e.g., insecure system accounts, open ports, missing patches) that had been identified by the Space and Naval Warfare Systems Center (SPAWAR). The two teams were tasked with bringing their system up to Fleet specification while avoiding changes that did not conform to Fleet regulations.

All Fleet ITs had unlimited access to the Internet. Navy/Marine Corps Internet (NMCI) files would be downloaded on request to the White Team. The Information Condition (INFOCON), based on the likelihood of cyber attacks, was initially set to INFOCON 5 (Red, Critical). After 120 minutes, it was lowered to INFOCON 3 (Yellow, Elevated). Scoring procedures were determined in advance, taking into account the severity of each vulnerability

and whether teams made corrections that were not authorized by the appropriate Program of Record.

Exercise procedures were introduced by the White Team in accord with a written script. When ACE/CID or Fleet IT team members identified a vulnerability, they were to determine how to correct it and then inform a White Team member of their findings before proceeding. Secondary issues that would take excessive time to execute were noted but then waved by the White Team. At the end of the testing, the Red Team assessed the team results by checking the system and the White Team notes to determine what vulnerabilities had been located and whether correct actions had been specified or performed to remove them.

Table 10 shows the points scored (out of a possible 105) by each of the four teams, along with their totals. As suggested by this table, the Fleet IT teams achieved higher scores on this exercise than the ACE/CID teams. This difference falls just short of statistical significant at $p < .05$.

Table 10. Security Testing Results

Topics	ACE/CID Teams				Fleet Teams			
	Team 1-1	Team 1-2	Team 2-1	Team 2-2	Team 1-1	Team 1-2	Team 2-1	Team 2-1
Required Patches	4	4	10	4	0	6	8	4
Symantec Mail Security for Exchange	0	0	0	0	0	0	0	0
Symantec Server Antivirus	1	1	1	1	2	7	8	8
Service Passwords	0	10	4	4	4	10	5	5
Banner Needs	0	0	0	0	0	0	0	0
Unauthorized Software	0	1	0	1	4	1	4	1
Domain User with Administrative Rights	0	2	0	0	0	0	0	0
INFOCON 3 Settings	2	12	9	6	13	18	17	14
Switch and Router Passwords (Bonus)	5	3	5	5	5	0	5	0
Team Total Points	12	33	29	21	28	42	47	32
Overall (420 possible)	ACE/CID Total = 95 (23%)				Fleet Total = 149 (35%)			

Note for Table 10: The numbers under the word “Team” in the column headings indicate the team and the cycle (i.e., Team 1-1 = Team 1/Cycle 1).

Security testing proved to be unexpectedly difficult for all four teams. The primary issue appeared to be the inability of the groups to execute as teams. Teams were reminded during Day 4 orientation of the importance of teamwork. They were told to organize and to pick their own team leader(s). The teams picked leaders, but they failed to organize as a unit. For instance, they often failed to identify and perform tasks that could be done in parallel. Some secondary

failures were also not anticipated. Some teams did not access their documentation, or, if they did, they did not read it carefully. Finally, because of time pressures, the one-on-one tutorial methodology used elsewhere in ACE/CID training was abandoned for a significant portion of the Security training.

d. Design and Implementation (Day 5)

A significant proportion of Navy ITs are deployed in non-Fleet, non-shipboard environments that require the design and implementation of an entire IT system. Day 5 focused on these issues. The IWAR-L participants were tested on their ability to operate in an environment to which they had never been exposed and to understand systems with which they had never worked.

Design and Implementation Procedure

Participating ITs were given a rack of (physical) equipment and a set of objectives that covered matters such as network redundancy, server redundancy, policy requirements, and traffic optimization. They were then allowed 7 hours to design, implement, and configure a new Information Technology system with network and server support that would achieve 24 objectives, which were selected for IWAR-L from various land-based assignment tasks.

The three classes of objectives were

1. Critical (e.g., establish a Windows domain; enable Internet access for all clients; automatically configure client machine Transmission Control Protocol/Internet Protocol (TCP/IP))
2. Secondary (e.g., create/test user accounts for key commanders; display a warning banner at logon; have workstations automatically lock after 5 minutes of inactivity)
3. Optimization (e.g., distribute server roles to maximize efficiency, enable DNS servers to resolve IP addresses to names).

In each of the two IWAR-L cycles, six ACE/CID students worked as a single team in one laboratory and six Fleet ITs worked as a single team in the other laboratory. Each group chose its own leader. For this exercise, White Team observers were present in each laboratory. They scored, with comments, the design and implementation produced. Achievement of each objective was rated on a 6-point scale, with 0 representing no accomplishment and 5 representing full and correct achievement of the objective. Points in between were anchored to specific implementation possibilities.

Design and Implementation Results

As suggested by Table 11 (Day 5 objectives successfully achieved) and Table 12 (Day 5 scores), neither group did well in this exercise—successfully achieving only about 60% of the objectives and receiving a little over half of the total possible score points. At the end of Cycle 1, the ACE/CID team had successfully achieved three more objectives than the Fleet IT team (see Table 11) and was slightly ahead in the scoring (see Table 12). At the end of Cycle 2, the Fleet IT team had successfully achieved four more objectives than the ACE/CID team (see Table 11) and came out ahead in the scoring (see Table 12). This difference falls just short of statistical significance at $p < .05$.

Table 11. Design and Implementation Objectives Scored 4–5 (Successful)

Objectives	Cycle 1		Cycle 2		Fleet IT Total Objectives Scored 4–5	ACE/CID Total Objectives Scored 4–5
	Fleet IT Objectives Scored 4–5	ACE/CID Objectives Scored 4–5	Fleet IT Objectives Scored 4–5	ACE/CID Objectives Scored 4–5		
Critical (7 objectives)	5	5	2	2	7	7
Secondary (12 objectives)	3	6	5	1	8	7
Optimization (3 objectives)	0	0	0	0	0	0
Totals	8	11	7	3	15	14

Table 12. Design and Implementation Scores (Day 5)

Objectives	Cycle 1		Cycle 2		Fleet Total Scores (Out of 220)	ACE/CID Total Scores (Out of 220)
	Fleet IT Total Scores (Out of 110)	ACE/CID Total Scores (Out of 110)	Fleet IT Total Scores (Out of 110)	ACE/CID Total Scores (Out of 110)		
Critical (7 objectives)	27.5	24.7	12	11	39.5	35.7
Secondary (12 objectives)	31.5	39.5	35	5	66.5	44.5
Optimization (3 objectives)	5.5	3.3	2	1	7.5	4.3
Totals	64.5	67.5	49	17	113.5	84.5

Even though the Cycle 2 Fleet IT team scored well with regard to the specific objectives and scoring rubric established for Day 5, White Team observers pointed out that the team had also damaged some system components beyond repair. The system should be “slicked,” and the

team should either start over or receive a score of zero for the day. The scoring rubric will have to be changed for Phase 2 IWAR to take account of such damage.

As in the Day 4 Security exercise, Day 5 tested the ability of the groups to organize. The groups could not achieve all the objectives by addressing them one at a time. The groups needed to assign different members to work on different objectives that could be accomplished in parallel. Both groups in both cycles had difficulty in organizing this way. The systematic approach taken by the ACE/CID students, which had served them well in earlier exercises, slowed their work and made a serial approach particularly troublesome.

During the off-site lunch period of Cycle 1, some members the Fleet IT group attempted to contact outside help. Members of the White Team successfully interrupted two of these attempts before any advice was received. A brief inquiry did not reveal any successful or other attempts. A closely supervised lunch was provided on-site during Cycle 2.

F. IWAR-D

1. IWAR-D Organization

IWAR-L provided a comprehensive, objective, and controlled comparison of ACE/CID students' knowledge and skills with those of experienced Fleet ITs. IWAR-D was intended to take the next step and assess the dockside performance of ACE/CID students using installed shipboard systems (to the extent permitted by shipboard conditions and available human observers).

Initially, IWAR-D was to compare the performance of ACE/CID students with Fleet ITs, in shipboard, dockside environments. However, given the results from IWAR-L and the constraints of shipboard IT activity, such a comparison appeared to deliver little additional information relative to its cost and effort. As subsequently planned with Navy coordination and assistance, IWAR-D was instead designed to observe only the performance of the ACE/CID students in shipboard, dockside environments.

2. IWAR-D Schedule

IWAR-D consisted of two 5-day cycles, 3–7 August and 10–14 August 2009. Cycle 1 assessed the shipboard performance of six ACE/CID students, and Cycle 2 assessed the performance of the remaining six ACE/CID students. Two ships, *USS Preble* (DDG 88) and *USS Milius* (DDG 69) (see Figure 3), were made available for both cycles. Three ACE/CID students were assigned to each ship in each cycle.




USS Preble (DDG 88)



USS Milius (DDG 69)

Figure 3. Ships Used During IWAR-D

Two primary sources of data were obtained from IWAR-D: a two-page survey completed by shipboard supervisors and a list or diary completed by the ACE/CID students. Figure 4 shows an example survey.



INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive / Alexandria, Virginia 22311-1882 / Telephone 703-845-2000

USS Preble (DDG-88) (Cycle 2)
Your opinions and impressions are critically important at this stage of IWAR testing. Please complete this survey at the end of each day this week. Thank you for your help!

Date _____

Preventive Maintenance

About how many Preventive Maintenance tasks did the DARPA/ITs work on today? _____

How difficult were the Preventive Maintenance tasks assigned to DARPA/ITs?

Very Easy	Easy	Average	Difficult	Very Difficult
-----------	------	---------	-----------	----------------

What percent of Preventive Maintenance tasks assigned to the DARPA/ITs did they complete successfully?

0-20%	20-40%	40-60%	60-80%	80-100%
-------	--------	--------	--------	---------

What were the DARPA/ITs' contributions in completing Preventive Maintenance tasks assigned to them?

None	Few/Minor	Some/Average	Many/Substantial	Major/Essential
------	-----------	--------------	------------------	-----------------

What was the value of Preventive Maintenance tasks completed by the DARPA/ITs to ship operations?

Small	Minor	Average	Substantial	Major
-------	-------	---------	-------------	-------

What amount of supervision was required for the DARPA/ITs' Preventive Maintenance task activities?

Constant	Frequent	Occasional	Rare	None
----------	----------	------------	------	------

Are there comments on the DARPA/ITs' Preventive Maintenance performance, value, and abilities you'd like to add? Please write them here and/or on the back of this sheet.

Scheduled Maintenance

About how many Scheduled Maintenance tasks did the DARPA/ITs work on today? _____

How difficult were the Scheduled Maintenance tasks assigned to DARPA/ITs?

Very Easy	Easy	Average	Difficult	Very Difficult
-----------	------	---------	-----------	----------------

What percent of Scheduled Maintenance tasks assigned to the DARPA/ITs was completed successfully?

0-20%	20-40%	40-60%	60-80%	80-100%
-------	--------	--------	--------	---------

What were the DARPA/ITs' contributions to completing the Scheduled Maintenance tasks assigned to them?

None	Few/Minor	Some/Average	Many/Substantial	Major/Essential
------	-----------	--------------	------------------	-----------------

What was the value of Scheduled Maintenance tasks completed by the DARPA/ITs to ship operations?

Small	Minor	Average	Substantial	Major
-------	-------	---------	-------------	-------

What amount of supervision was required for the DARPA/ITs' Scheduled Maintenance task activities?

Constant	Frequent	Occasional	Rare	None
----------	----------	------------	------	------

Are there comments on the DARPA/ITs' Scheduled Maintenance performance, value, and abilities you'd like to add? Please write them here and/or on the back of this sheet.

Trouble Ticket Responses

About how many Trouble Tickets did the DARPA/ITs work on today? _____

How difficult were the Trouble Ticket tasks assigned to DARPA/ITs?

Very Easy	Easy	Average	Difficult	Very Difficult
-----------	------	---------	-----------	----------------

What percentage of Trouble Ticket tasks assigned to the DARPA/ITs was completed successfully?

0-20%	20-40%	40-60%	60-80%	80-100%
-------	--------	--------	--------	---------

What were the DARPA/ITs' contributions to completing the Trouble Ticket tasks assigned to them?

None	Few/Minor	Some/Average	Many/Substantial	Major/Essential
------	-----------	--------------	------------------	-----------------

What was the value of Trouble Ticket tasks completed by the DARPA/ITs to ship operations?

Small	Minor	Average	Substantial	Major
-------	-------	---------	-------------	-------

What amount of supervision was required for the DARPA/ITs' Trouble Ticket task activities?

Constant	Frequent	Occasional	Rare	None
----------	----------	------------	------	------

Are there comments on the DARPA/ITs' Trouble Ticket performance, value, and abilities you'd like to add? Please write them here and/or on the back of this sheet.

Emerging Problems

About how many Emerging Problems did the DARPA/ITs work on today? _____

How difficult were the Emerging Problem tasks assigned to DARPA/ITs?

Very Easy	Easy	Average	Difficult	Very Difficult
-----------	------	---------	-----------	----------------

What percentage of Emerging Problem tasks assigned to the DARPA/ITs was completed successfully?

0-20%	20-40%	40-60%	60-80%	80-100%
-------	--------	--------	--------	---------

What were the DARPA/ITs' contributions to completing the Emerging Problem tasks assigned to them?

None	Few/Minor	Some/Average	Many/Substantial	Major/Essential
------	-----------	--------------	------------------	-----------------

What was the value of Emerging Problem tasks completed by the DARPA/ITs to ship operations?

Small	Minor	Average	Substantial	Major
-------	-------	---------	-------------	-------

What amount of supervision was required for the DARPA/ITs' Emerging Problem task activities?

Constant	Frequent	Occasional	Rare	None
----------	----------	------------	------	------

Are there comments on the DARPA/ITs' Emerging Problem performance, value, and abilities you'd like to add? Please write them here and/or on the back of this sheet.

Other Activity (please identify) _____

Figure 4. Example IWAR-D Survey

The survey consisted of four main sections: Preventive Maintenance, Scheduled Maintenance, Trouble Ticket Responses, and Emerging Problems. The Emerging Problems tasks were

those found while correcting other problems. The survey included a final, free form section called Other Activity, which could be used to discuss tasks that did not fit into the other four categories.

The diaries consisted of a straightforward list of tasks attempted and the actions taken to accomplish them. In general, the diaries appear to provide a more accurate description of the IWAR-D activities and results than the surveys provided. For instance, one ship’s set of survey responses reports 58 tasks for Cycle 2, with exactly the same number of tasks in each of the four task categories in 3 of the 4 days reported. These survey responses do not match the ACE/CID student diary reports for that week, which list a total of 8 tasks assigned that week and is precise in describing each task and the steps taken to accomplish it. Surveys from that ship for its other cycle report also do not match the tasks assigned as described in the respective task diaries.

Diary and survey reports from the other ship differ much less. These reports and the data reported here are based on survey data from that ship and the task diaries from both ships.

3. IWAR-D Results

Overall, the ACE/CID students completed about 38 tasks during IWAR-D—16 in Cycle 1 and 22 in Cycle 2. The number of tasks completed was affected by shipboard schedules and activities. For example, no tasks were performed for 2 days aboard the *Preble* in Cycle 1 because of prescheduled drills and exercises. The categories assigned to tasks differed considerably between the ship surveys used here and the task diaries, so all tasks are reported together. However, it does appear that most of the tasks were either trouble ticket responses or emerging problem tasks. Very few preventative maintenance or scheduled maintenance tasks were assigned to the ACE/CID students. Assigning the values 1–5 to each survey dimension (task difficulty, tasks successfully completed, student contributions to completion, value of task completion to ship’s operations, and amount of supervision required) lead to the results shown in Table 13.

Table 13. Survey Responses Concerning Tasks Assigned to ACE/CID Students

Dimension	Average	Comment
Task difficulty	3.6	Above average
Tasks successfully completed	4.3	About 87% successfully completed
Student contributions to completions	3.8	Above average
Value of task completion to ship’s operations	3.9	Above average
Amount of supervision required	3.6	Occasional to rare

The task diaries and the surveys do not indicate whether the ACE/CID students were assigned any tasks that they were not prepared to solve. Some tasks were not solved because the ACE/CID students were not permitted access to a component or because the time allocated for the day's IWAR-D testing ran out.

Most of the ACE/CID students' contributions to their tasks were rated as substantial or essential. In two instances, the ACE/CID students solved problems that the Fleet ITs had not been able to solve either with their own solutions or with solutions they had subsequently received from SPAWAR.

The ACE/CID students solved about half of their tasks with little or no supervision. Supervision of the ACE/CID students typically diminished in the IWAR-D cycles after the shipboard ITs had an opportunity to observe their performance.

G. IWAR-A

1. IWAR-A Organization

IWAR-A was the final stage of Phase 1. It was intended to assess the performance of ACE/CID students deployed aboard ships at sea. As with IWAR-D, testing was to use shipboard systems and employ human observers to the extent permitted by both shipboard and deployment conditions.

2. IWAR-A Schedule

Again, because of the wealth of data provided by IWAR-L, likely additional information to be obtained, and constraints of time and resources, IWAR-A was limited to observations of three ACE/CID students who served aboard the *USS Vandegrift* (FFG-48) (see Figure 5) on 17–20 August 2009. The students selected for IWAR-A were the three whose next duty stations were shore based.



***USS Vandegrift* (FFG-48)**

Figure 5. Ship Used During IWAR-A

3. IWAR-A Results

IWAR-A data consists of the task diaries kept by the ACE/CID students and a narrative report from the Senior Chief who served as their monitor. From the start, the ACE/CID students were assigned system-wide tasks that are typically addressed by senior ITs who have taken pre-requisite “C” school courses and have several years of fleet experience.

On the first day, the students made registry changes on several servers to install anti-virus software on the ship’s Secret Internet Protocol Router Network (SIPRNET), reloaded a router that was not routing but was not reporting any problems, and added the Exchange address into the registry key. These actions were essential for successful operation of the ship’s Information Technology systems. If any were done incorrectly, the results would have been debilitating.

On the second day, the ACE/CID students discovered a potential “emerging problem”—one of back-up systems was not functional—that had not been reported. The problem was solved when the ACE/CID students installed the correct software on the system and informed the rest of the system of its availability, when and if needed.

The work of the ACE/CID students continued in the third and fourth day as regular members of the ship’s IT company.

Despite being in a new and unfamiliar environment, the three ACE/CID students appear to have followed the careful and systematic processes they had been taught. Evidence in the task diary showed that they (1) checked to see what the core problem was and whether it matched the symptoms reported, (2) corrected the problem without making actual or potentially harmful changes in the system, and (3) ensured that their work correctly solved the problem.

All three ACE/CID students participated equally in technical activities. The success of their brief tour was not a matter of a single superior student solving problems for all three.

H. Discussion

The IWAR-L results suggest that the ACE/CID training achieved its goal: to demonstrate that the content and methodologies planned for the computerized DT could achieve the DARPA objective of producing, in 16 weeks, new sailors who have the knowledge and skills of experienced Fleet ITs.

In the Knowledge Testing and Troubleshooting exercises, the results appear to have exceeded DARPA’s goal of showing parity between the two groups. The ACE/CID students were superior in these activities. The Fleet ITs came out ahead in the Day 4 Security scores, although the difference fell short of statistical significance. This result may reflect the fact that the one-on-one instructional methodology used elsewhere in ACE/CID training was abandoned

for portions of the Security training. The Fleet ITs also scored higher in the Day 5 Design and Implementation exercises, but these scores again fell short of statistical significance. The White Team expressed reservations about the Fleet IT scores because of damages they made to the systems in implementing IWAR objectives. The scoring rubric for Day 5 will need to be changed for Phase 2 IWAR.

As stated, the DARPA Education Dominance Program is a research effort. Much is being and will be learned as it progresses, which seems appropriate for a program that has ambitious goals of considerable consequence to the Navy. Testing programs as elaborate and extensive as IWAR-L are usually preceded by careful pilot testing of materials and procedures. Between the pace of development and the need to provide data that would reliably inform decisions concerning Phase 2, neither resources nor, especially, time was available for pilot testing. The organizers and conductors of IWAR-L are to be congratulated for ensuring that it operated as smoothly as it did.

Overall, the ACE/CID training appears to have more than achieved its objectives. Using the content and methodology to be incorporated into the DT, ACE/CID training produced students who demonstrated equal—and, in many cases, superior—knowledge and skill compared to that of experienced Fleet ITs.

Phase 2 of the DARPA Education Dominance Program will show whether ACE/CID content and methodology can be digitized for the DT and then prove to be as successful as it was in IWAR Phase 1. Success in Phase 2 would hold considerable promise not just for Navy Information Technology training, but also for large portions of Department of Defense (DoD) specialized skill training. The superior performance of ACE/CID students over Fleet ITs on the 21 DT-related IWAR Knowledge Test items (see Table 7) suggests that the DARPA Educational Dominance Program—if successful in Phase 2—should advance the state of training art and practice and contribute significantly to Navy and DoD operational capabilities.

References

- Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher* 13, 4–16.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Coulson, J. E. (Ed.) (1962). *Programmed learning and computer-based instruction*. New York, NY: John Wiley.
- Fletcher, J. D. (2001). What do sharable instructional objects have to do with intelligent tutoring systems, and vice versa? *International Journal of Cognitive Ergonomics* 5, 317–333.
- Fletcher, J. D. (2009). Education and training technology in the military. *Science* 323, 72–75.
- Fletcher, J. D., and Chatham, R. E. 2010. Measuring return on investment in military training and human performance. In J. Cohn & P. O'Connor (Eds.), *Human performance enhancements in high-risk environments* (pp. 106–128). Santa Barbara, CA: Praeger/ABC-CLIO.
- Fletcher, J. D., and Rockway, M. R. (1986). Computer-based training in the military. In J. A. Ellis (Ed.), *Military contributions to instructional technology* (pp. 171–222). New York, NY: Praeger Publishers.
- Ford, J. D., Slough, D. A., and Hurlock, R. E. (1972). *Computer-assisted instruction in Navy technical training using a small dedicated computer system: Final report*. Research Report No. SRR 73-13. San Diego, CA: Navy Personnel Research and Development Center.
- Graesser, A. C., and Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal* 31, 104–137.
- Graesser, A. C., D'Mello, S. K., and Cade, W. (In Press). Instruction based on tutoring. In R.E. Mayer and P.A. Alexander (Eds.), *Handbook of research on learning and instruction*. New York: Routledge Press.
- Luckin, R., Koedinger, K. R., and Greer, J. (Eds.). (2007). *Artificial intelligence in education*. Amsterdam: IOS Press.
- Potka, J., Massey, L. D., and Mutter, S. A. (Eds.) (1988). *Intelligent tutoring systems: Lessons learned*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Scriven, M. (1975). Problems and prospects for individualization. In H. Talmage (Ed.), *Systems of individualized education* (pp. 199–210). Berkeley, CA: McCutchan.
- Sleeman, D., and Brown, J. S. (Eds.) (1982). *Intelligent tutoring systems*. New York, NY: Academic Press.

Glossary

ACE	Advanced Concepts in Education
AFQT	Armed Forces Qualification Test
ANA	Advanced Network Analyst
CASREP	Casualty Report
CID	Center for Information Dominance
DARPA	Defense Advanced Research Projects Agency
DDG	guided missile destroyer
DNS	Domain Name System
DoD	Department of Defense
DT	Digital Tutor
FFG	guided missile frigate
FSET	Fleet Systems Engineering Team
FY	Fiscal Year
IDA	Institute for Defense Analyses
INFOCON	Information Condition
IOS	Internetwork Operating System
ISSM	Information Systems Security Manager
IT	Information Systems Technician
IWAR	Information Warfare
IWAR-A	IWAR-At Sea
IWAR-D	IWAR-Dockside
IWAR-L	IWAR-Laboratory
JNETCORE	Journeyman-Networking Core
NMCI	Navy/Marine Corps Internet
NSVT	Network Security Vulnerability Technician
NTCSS	Navy Tactical Command Support System
OSPF	Open Shortest Path First
R&D	research and development

RF	radio frequency
SIPRNET	Secret Internet Protocol Router Network
SME	subject matter expert
SPAWAR	Space and Naval Warfare Systems Center
TCP/IP	Transmission Control Protocol/Internet Protocol

REPORT DOCUMENTATION PAGE*Form Approved*
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE January 2010		2. REPORT TYPE Final		3. DATES COVERED (From–To) April 2009 – November 2010	
4. TITLE AND SUBTITLE Phase 1 IWAR Test Results				5a. CONTRACT NUMBER DASW01 04 C 0003	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) J.D. Fletcher				5d. PROJECT NUMBER	
				5e. TASK NUMBER DA-2-2896	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Institute for Defense Analyses 4850 Mark Center Drive Alexandria, VA 22311-1882				8. PERFORMING ORGANIZATION REPORT NUMBER IDA Document D-4047 Log: H10-000220	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Defense Advanced Research Projects Agency Defense Sciences Office 3701 North Fairfax Drive Arlington, VA 22203-1714				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited. (14 April 2010)					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT In Phase 1 of DARPA's Information Warfare (IWAR) project, 12 Information Systems Technician (IT) "A" School candidates from the Navy's Center for Information Dominance (CID) were trained in a 16-week Advanced Concepts in Education (ACE) program of instruction. Except for about 2 weeks when early available components of the computerized Digital Tutor (DT) were used, highly qualified human tutors conducted this training. The ACE/CID program was assessed by comparing the IT knowledge and skills of its students with those of 12 Fleet ITs with 3–12 years of IT experience. Testing involved 4 hours of knowledge testing, 13.25 hours of troubleshooting, 3.5 hours of security performance, and 7 hours of design and implementation. ACE/CID students scored higher than Fleet ITs on 11 of the 12 knowledge test topics, solved 99 troubleshooting problems compared to 79 solved by Fleet ITs, made fewer harmful changes (8 vs. 18), and verified both more problems (97% vs. 85%) and more solutions (95% vs. 77%). Additional shipboard testing involved ACE/CID students alone. In dockside testing, they solved 87% of assigned tasks. Unsolved tasks involved lack of time or required access to prohibited IT components. ACE/CID students at sea successfully performed system-wide tasks that are typically assigned to senior ITs.					
15. SUBJECT TERMS education, intelligent tutoring systems, Information Systems Technician (IT), technician training, technology-based instruction, training, tutorial instruction					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT Uncl.	b. ABSTRACT Uncl.	c. THIS PAGE Uncl.			LCDR Joseph Cohn
			SAR	39	19b. TELEPHONE NUMBER (include area code) 571-830-4664