Evaluating Training Effectiveness: Instructional Support Software from Squads to Schoolhouses

Sae Schatz
UCF Institute for Simulation & Training
Orlando, FL
sschatz@ist.ucf.edu

Roberto Champney
Design Interactive, Inc.
Orlando, FL
roberto@designinteractive.net

Stephanie Lackey
UCF Institute for Simulation & Training
Orlando, FL
slackey@ist.ucf.edu

Cynthia Oakes, Robert Dunne
UCF Institute for Simulation & Training
Orlando, FL
coakes@ist.ucf.edu
rdunne@ist.ucf.edu

ABSTRACT

As instructional simulations become more prolific, both official and unofficial instructors are expected to facilitate their use in training and education. For formally trained and billeted instructors, this entails the expected instructional duties as well as often-challenging technology administration. Informal instructors (such as small unit leaders) are expected to incorporate simulation into their units' training activities, too; however, small unit leaders typically receive minimal guidance on how to effectively facilitate simulation-based training. Instead, when it comes to implementing simulations, many small unit leaders may "wing it," rather than carefully planning, executing, and evaluating the training in accordance with instructional best practices.

To address these concerns, the project team developed an Instructional Support System (ISS) that, in part, helps instructors (across all experience levels) design and deliver more effective and efficient simulation-based training. The ISS software integrates with the Deployable Virtual Training Environment (DVTE), a laptop-based simulation platform used throughout the Marine Corps. The ISS facilitates numerous instructional tasks, from trainee management to after-action review, in addition to providing resources that help users develop more systematic, instructionally sound curricula.

A two-fold investigation was conducted to assess the operational support capabilities of the ISS and to determine whether the ISS's lesson development support enabled small unit leaders to create more effective lesson plans. First, Marines (N = 57), trained in DVTE, were asked to perform two key tasks (i.e., launching a scenario and identifying a scenario for a given training objective) using either DVTE alone or DVTE and the ISS. We examined their efficiency and effectiveness to assess the systems' operational utility. Second, Marine Sergeants (N=80) at the Enlisted Professional Military Education (E-PME) schoolhouse were asked to develop lessons on Call For Fire. Using control and experimental conditions, we compared the utility, effectiveness, and appeal of the ISS against DVTE alone (i.e., the status quo). In both cases, the participants using the ISS significantly outperformed those in the DVTE-only conditions.

ABOUT THE AUTHORS

Dr. Sae Schatz is an Assistant Professor with the University of Central Florida Institute for Simulation & Training (UCF-IST), and she is one of the lead scientists on the Next Generation Expeditionary Warfare – Intelligent Training (NEW-IT) contract, which sought to improve Marine Corps simulation-based training.

Dr. Roberto Champney is a Senior Research Associate at Design Interactive, Inc. and is currently supporting training system design and evaluation for the Office of Naval Research's Human Performance, Training and Education (HPT&E) Program. His work focuses on field analysis, design and evaluation of interactive systems.

Dr. Stephanie Lackey is the Director of the Applied Cognition & Training in Immersive Virtual Environments (ACTIVE) Lab at UCF-IST.

Cynthia Oakes is an undergraduate psychology student at UCF-IST. She recently completed her undergraduate honors thesis on social group affiliation and self-esteem.

Robert Dunne is a doctoral candidate in Instructional Technology and graduate research assistant at UCF-IST. His latest research involves the operationalization of scenario complexity and motivation in serious games.

Evaluating Training Effectiveness: Instructional Support Software from Squads to Schoolhouses

Sae Schatz
UCF Institute for Simulation & Training
Orlando, FL
sschatz@ist.ucf.edu

Roberto Champney
Design Interactive, Inc.
Orlando, FL
roberto@designinteractive.net

Stephanie Lackey
UCF Institute for Simulation & Training
Orlando, FL
slackey@ist.ucf.edu

Cynthia Oakes, Robert Dunne
UCF Institute for Simulation & Training
Orlando, FL
coakes@ist.ucf.edu
rdunne@ist.ucf.edu

DEMAND SIGNAL

It has been a decade since the terrorist attacks of 9/11, and, by now, the criticality of distributed and full-spectrum operations, and the complementary notion of *strategic corporals* (Krulak, 1999), have become prevailing concepts. As a result, emphasis continues to be placed on cognitive readiness, that is, the idea that personnel across all echelons must possess the "mental preparation (including skills, knowledge, abilities, motivations, and personal dispositions)…to establish and sustain competent performance in the complex and unpredictable environment of modern military operations" (Morrison & Fletcher, 2002, ES-1).

The Marines are particularly focused on this challenge. General James Amos, 35th Commandant of the Marine Corps, listed training and education among his four priorities: "We will better educate and train our Marines to succeed in distributed operations and increasingly complex environments. We will invest more in the education of our NCOs and junior officers, as they have assumed vastly greater responsibilities in both combat and garrison" (Amos, 2010, p. 8).

To meet these aims the Marines require instructional technologies that facilitate the training and education of higher-order cognitive skills—but that do so in a flexible, efficient, and sustainable manner (see Becker & Schatz, 2010 for a detailed discussion). Simulation-based training systems can meet this need. However, most real-world instructional simulations suffer from logistical problems that limit their practical usability, efficacy, and deployability.

A key limitation of typical simulation-based training systems is the extensive instructor participation required to plan, execute, and troubleshoot their use (Salas et al., 2000; Ross et al., 2005; Oser et al., 1999). At best, this places high resource demands on instructors, and at worse, in deployed settings where formal instructional personnel are unavailable, this may lead to less-than-optimal training outcomes (Loftin et al., 2004).

In this paper, we describe our efforts to design and evaluate an Instructional Support System (ISS) for USMC simulation-based training. This system was developed to enhance the instructional effectiveness and operational utility of training simulations. That is, the ISS was designed to help formal Marine instructors do their jobs more efficiently and small unit leaders complete their training duties more effectively.

Official and Unofficial Instructors

The Marine Corps has a tradition of high-quality instructor preparation. For instance, Marine combat instructors attend an intensive two-month course on leadership, subject matter, regulations, and instructional techniques. This helps them develop the knowledge and skills necessary to formally train other personnel. Following completion of this course, new combat instructors additionally receive a year of apprenticeship, under the tutelage of experienced trainers, before they lead their own classes (Combat Instructor School Overview, n.d.). We refer to such personnel as "official instructors" because they hold instructor billets and receive formal training on instructional best practices.

However, just as "every Marine is a rifleman," every Marine is also expected to be an instructor, in some capacity. Outside of the schoolhouses, a large cadre of unofficial instructors helps train, educate, and sustain the Corps' knowledge, skills, and attitudes. Some of these personnel receive limited training on how to be effective instructors, but others receive no formal instructional preparation at all. In this paper, we call these personnel "unofficial instructors" or "lay-instructors" because they are expected to facilitate instruction but lack extensive training on instructional best practices.

Simulation-Based Training and Education

As instructional simulations become more prolific, both official and unofficial instructors are expected to facilitate their use in training and education. For official instructors, this entails the expected instructional duties as well as simulation administration (e.g., loading scenarios, managing digital trainee records, "puckstering" simulated entities)—which can be slow and burdensome. These duties consume valuable time, taxing the scarce resources of qualified instructors.

Unofficial instructors are similarly expected to support the training, education, and technology use associated with instructional simulations. However, because these lay-instructors typically receive minimal guidance on how to effectively plan or execute simulation-based training events, they may "wing it," without using structured instructional plans. This can lead to less-than optimal training experiences (e.g., Loftin et al., 2004).

Section Summary

As this section described, the effectiveness and efficiency of military simulation-based training could be enhanced; however, current technologies suffer from practical limitations. For official instructors, simulation systems can be onerous to use—forcing them to focus valuable energy on "button-ology" rather than training content. For unofficial instructors, simulations may also be difficult to employ effectively, because they have little training or education in instructional best practices.

INSTRUCTIONAL SUPPORT SYSTEM

To help address the needs of both official and unofficial simulation instructors, our team designed the ISS and developed it under the Next Generation Expeditionary Warfare Intelligent Training (NEW-IT) project, a three-year Office of Naval Research (ONR) initiative that began in late 2008. NEW-IT's mandate was to deliver science and technology solutions that improve the effectiveness and efficiency of USMC simulation-based training, particular for higher-order cognitive skills instruction. (For additional programmatic information, including a list of the full research and development team, see the Schatz, Lackey, Stanney, & Schaffer, 2011 technical report).

The NEW-IT ISS is a software program that facilitates the planning and executing of simulation-based instruction (see Figure 1). Conceptually, it "wraps around" an existing simulation platform to help guide instructional preparation (e.g., lesson creation, instructional strategy selection, trainee record maintenance) and more easily facilitate instructional execution (e.g., monitoring trainees' status, collecting performance metrics, delivering after action reviews).

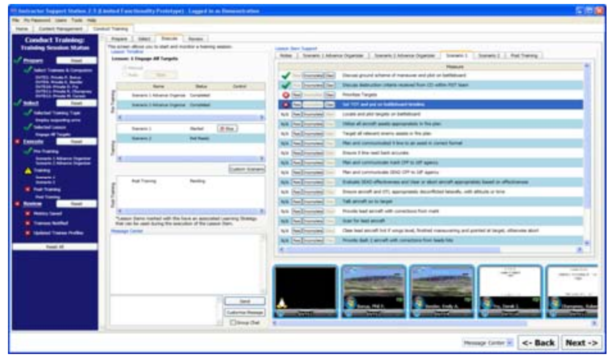


Figure 1. Screenshot of the ISS, during a training event

The initial ISS prototype was developed to support the Deployable Virtual Training Environment (DVTE) simulation, a multi-user laptop-based suite that the Marines fielded world-wide. However, the ISS software and, more importantly, its underlying research-based design were intended to be simulation agnostic.

ISS EFFECTIVENESS EVALUATION

The NEW-IT team conducted a two-pronged effectiveness evaluation of the ISS, focusing on both *operational effectiveness* (i.e., utility) and *instructional effectiveness* (i.e., facilitation of enhanced training). The operational effectiveness investigation (described as Study 1) emphasizes usability and efficiency enhancements to simulation-based training. This study addresses official instructors' needs, as described above. The instructional effectiveness investigation (described as Study 2) examines whether the ISS helps lay-instructors more effectively employ instructional planning and execution best practices.

STUDY 1: OPERATIONAL EFFECTIVENESS

To evaluate the operational effectiveness of the ISS, Marines and relevant USMC contractors were asked to complete two simulation instructor tasks, either with the ISS or with the DVTE simulation alone. The two activities, described in more detail below, involved (1) launching a training scenario and (2) selecting a training scenario that supports a given training objective. These two tasks were identified as challenging-yet-critical through field analyses and interviews with Marine stakeholders.

Participants

Fifty-seven (N=57) Marines and USMC contractors (i.e., simulation analysts and schoolhouse instructors) were recruited to participate in the study. All participants were screened for familiarity with the DVTE system, and only those competent with its use (e.g., had completed the DVTE Train the Trainer course) were invited to continue. Participants were assigned to one of five groups, based upon which ISS prototype version was available at the time (i.e., DVTE alone, ISS V1.0, ISS V1.5, ISS V2.0, ISS V2.5). The "DVTE alone" condition represents the operational status quo, and we refer to this condition as the baseline. Data collection events occurred every six months, between 2009 and 2011, beginning with baseline data collection and then focusing on newer versions of the ISS as iterative prototypes were developed. Ten participants per group were targeted yet due to availability of suitable participants, the actual population of each cohort ranged between 9 and 16.

Materials

The study utilized six DVTE computers, which were set-up so as to simulate a real-world USMC simulation-based training event (see Figure 2). That is, one computer ran the semiautonomous forces, and the other five represented trainee terminals. For the ISS versions of the task, one trainee terminal doubled as the instructor terminal (this has no impact on the task metrics which focused on preparation and set-up to execute training).

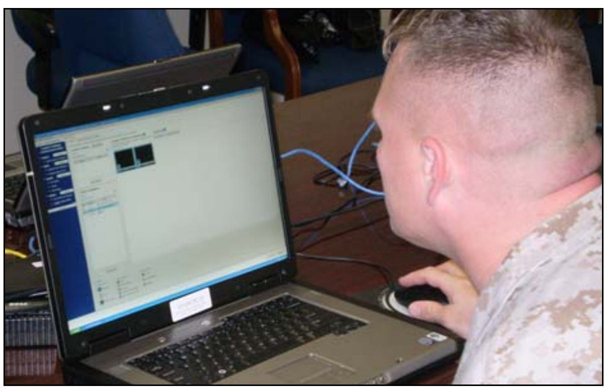


Figure 2. Participant using the ISS V2.0

Method

The study employed a quasi-experimental design with five groups. Each group's performance on the two tasks was compared. Participants in the first group completed the tasks using the DVTE simulation alone (i.e., baseline condition). Participants in the second, third, fourth, and fifth cohorts used DVTE plus ISS V1.0, V1.5, V2.0, and V2.5, respectively.

Task 1: Launch a Training Scenario
The first task involved participants launching a training scenario for a team of five imaginary trainees on the respective DVTE laptops. This involved configuring the simulation system and actually launching the scenario on all stations. Performance was measured in seconds; that is, how long it took to launch the scenario.

Task 2: Select an Appropriate Scenario
The second task involved participants identifying a scenario that suitably matched a given training objective. Performance was measured by evaluating whether the participant selected the "correct" scenario. The "correct" scenario was determined with support of SMEs who identified a scenario as the most suitable for the training objective selected for the experimental task.

Results

For the first task, the average time required to set-up a DVTE training scenario was compared across conditions. For the sake of readability, these figures have been converted into percentage scores, i.e., percent improvement as compared with baseline. Ultimately, participants using ISS V2.5 (the latest version of the ISS tested) were able to configure and launch a DVTE training scenario 79% more rapidly than those participants who used DVTE alone, which required over five-minutes to initiate (see Figure 3). Additionally, participants using ISS V2.5 selected the appropriate training scenario 98% more often than participants who used DVTE alone (see Figure 4).

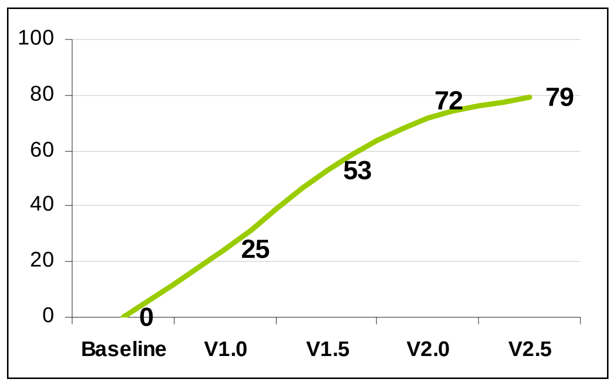


Figure 3. Percent improvement in setup time

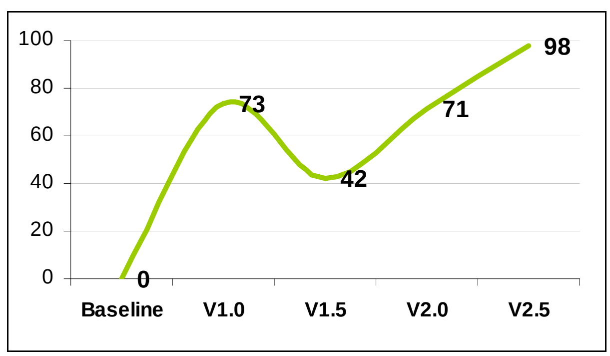


Figure 4. Percent improvement in scenario selection

These findings suggest that the ISS has indeed achieved observable gains in both operational efficiency and effectiveness. That is, the ISS successfully supports formal instructors by helping them more efficiently manage the simulation technology.

STUDY 2: INSTRUCTIONAL EFFECTIVENESS

Like other modes of instruction, simulation-based training and education are most effective when lessons are carefully planned and executed in accordance with instructional best practices. By instructional best practices, we mean adherence to Instructional Systems Design (ISD) principles, such as deliberate creation of lesson plans based on targeted instructional strategies.

Among its capabilities, the ISS provides resources to help instructors (particularly lay-instructors) more easily develop systematic, instructionally sound lessons. For our purposes, a "lesson" involves pre-practice instruction, a practical application, and post-practice feedback. Each ISS-based lesson is created for a specific topic and at a certain complexity level. Lessons are developed from preexisting lesson elements, such as documents, multimedia, metrics, instructional strategies, and simulation scenarios. Consider the example ISS-based lesson described in Table 1. (Once the ISS is transitioned into operational use, official instructors at schoolhouses will likely create these lesson items; however, the ISS allows all users to create or import lesson elements.)

Table 1. Example ISS Lesson Characteristics

Topic	Employ Supporting Arms
Complexity Level	Crawl (Novice)
Pre-Practice Materials *(i.e., Preparation)*	• Video of a fire team leader explaining key aspects of the task • Preparatory compare-and-contrast instructional exercise
Practice Materials *(i.e., Learning Practice Phase)*	• DVTE scenario designed for novice-level Employ Supporting Arms practice • Associated performance metrics
Post-Practice Materials *(i.e., Reflection)*	• Guided mission debrief worksheet and team feedback exercise

Instructional Best Practices

The ISS promotes several instructional best practices. For example, it encourages, but does not force, the use of pre-, during-, and post-practice interventions. It explicitly ties scenarios to training and readiness standards, and links additional performance metrics to lesson topics (making assessment more actionable than what is generally available simply through Training and Readiness Manuals). It also promotes adaptive (i.e., personalized) training by making lesson-selection recommendations based upon characteristics of the specific trainees, and it facilitates the learning and use of key instructional strategies.

Instructional Strategy Browser

The ISS includes an instructional strategy browser that lists specific instructional strategies available within the system. The strategies are categorized by trainee level (novice or advanced) as well as by training stage (pre, during, and post-practice). Each strategy is briefly described, so that lay-instructors will understand its purpose, rationale, and usage. A breakdown of the strategies included in the ISS is shown in Table 3. These particular strategies were selected because they are theorized to effectively support scenario-based instructional simulations. The efficacy of the strategies and positive learning impacts of their employment was previously evaluated in two laboratory-based experiments (see Fowlkes, Stagl, Schatz, & Lum, under review; Vogel-Walcutt, Marshall, Fowlkes, Schatz, Dolletski-Lazar & Nicholson, 2011).

In a previous empirical study, we demonstrated that lessons created according to the guidance of the ISS—specifically those that included one of the instructional strategies—prompt better trainee performance than lessons created without an intentional instructional strategy (Vogel-Walcutt et al., 2011). For this study, we sought to discover whether the ISS effectively guides lay-instructors to design more effective lessons and systematically employ ISS-suggested instructional strategies.

Method

We hypothesized that lay-instructors who used the ISS V2.5 would be more likely to design effective lessons. To test our hypothesis we conducted an empirical study with control and experimental conditions. Participants in both conditions were asked to develop a simulation-based training lesson for a given, fictional situation (described in more detail, below). All participants were able to access the DVTE simulation and its associated user manuals. Participants in the experimental condition were additionally able to access the ISS software (but no ISS user manual, since we believed that users, in ecological settings, would be unlikely to seek out user documentation and we wanted to avoid any positive bias in favor of the experimental group). A diagram of the experimental design is shown in Table 2.

Participants

As mentioned earlier, lay-instructors are often tasked with assembling lessons that involve instructional simulations. To allow for a realistic study, we sought a study population of unofficial simulation instructors. Specifically, eighty Marine sergeants (N = 80; male = 76; mean age = 25.68) were recruited from the Enlisted Professional Military Education (E-PME) schoolhouse in Quantico. These personnel spent an average of 6.38 years in the Marine Corps, and 65 of them had been deployed overseas. All participants were trainees in the E-PME Sergeants Course, which involves training on Call For Fire operations and DVTE simulation usage.

Table 2. Lay-Instructor Experimental Design Diagram

	O	X	O
Experimental	Demographics Call For Fire Knowledge Pretest	Create a Call For Fire training plan *with* ISS Software (1 hour)	Workload Reactions Instructional knowledge posttest
Control		Create a Call For Fire training plan *without* ISS Software (1 hour)	

Participants were divided into four, 20-person cohorts. These cohorts were established by E-PME for logistical reasons and were not an intentional part of the study design. Two cohorts were assigned to the experimental condition, and two were assigned to the control condition. All participants in a cohort completed the experiment concurrently.

Experimental Task

All participants were asked to create a training plan for a fictional situation, either with or without the aid of the ISS. The directions stated:

> Imagine that you have been asked to teach a class on Call For Fire (CFF). Your class is three-hours long, and it must include some scenario-based training. In other words, you need to use the DVTE simulator as part of the lesson. Your trainees are fresh USMC recruits (mainly PFCs), but they deploy in about two months—so you need to make sure they know the material.

The directions went on to list eight specific learning objectives that the fictional PFC trainees needed to learn, such as “the six elements of a CFF transmission.” Although the participants were not informed of this, two learning objectives involved declarative knowledge, two were procedural, two were conceptual, and two involved generalizability of knowledge.

To replicate realistic time constraints, participants were allowed only one hour to construct their lesson plans. Also, to replicate realistic resource constraints, trainees were not provided with technical support or an ISS user manual. Finally, to make the two conditions more comparable, all participants developed their lesson plans on paper. (The DVTE simulation system does not readily facilitate software-based creation of lesson plans

Table 3. Instructional Strategies, and their Characteristics, Included in the ISS

Strategy Name (as listed in the ISS)	Pre	During	Post	Novice	Advanced
Compare and Contrast	✓	✓	✓	✓	✓
Create Disequilibrium		✓			✓
Demonstrate	✓			✓	
Design Event-Based Training		✓		✓	✓
Direct Attention	✓			✓	✓
Encourage Expertise	✓			✓	
Encourage Self-Correction			✓	✓	
Link to other Learning and Practice			✓	✓	
Mentally Rehearse	✓			✓	✓
Metacognitive	✓	✓	✓	✓	✓
Provide Complexity		✓			✓
Trainee Misconceptions	✓		✓	✓	✓
Use Case-Based Self-Critique			✓		✓
Use Stories to Address the Messiness of Real Life			✓		✓
Vary Practice		✓			✓
What-If			✓		✓

nor does it include a library of lesson elements, even though these features are part of the ISS.)

Apparatus

Demographics

The demographics questionnaire included general questions (e.g., gender and age), as well as more targeted questions, such as how long each respondent had served in the Marines and whether respondents had had previous experience as an instructor.

Call For Fire Pretest

In an effort to identify possible mediating factors, participants were given a 19-item multiple choice pretest to assess their declarative, procedural, and conceptual knowledge of Call For Fire operations.

Lesson Plan Creation Template

As mentioned above, all participants completed Call For Fire lesson plans on paper. Participants received a lesson plan template worksheet with which to organize their responses. The template included columns labeled: item title, time, training phase, material (optional), instructor activity, learning objective, and instructional strategy. Additionally, one line of the worksheet was already filled-out, as an example. See Figure 5.

It should be noted that the template prompted trainees to design more methodical lesson plans than they would likely otherwise create. This may have given control-group participants, in particular, an extra advantage that they would not receive in operational settings; however, we felt that it was necessary to give the control group a fair chance to provide the specific types of information that are appropriate for ISD-informed lesson planning.

Subjective Workload Measure

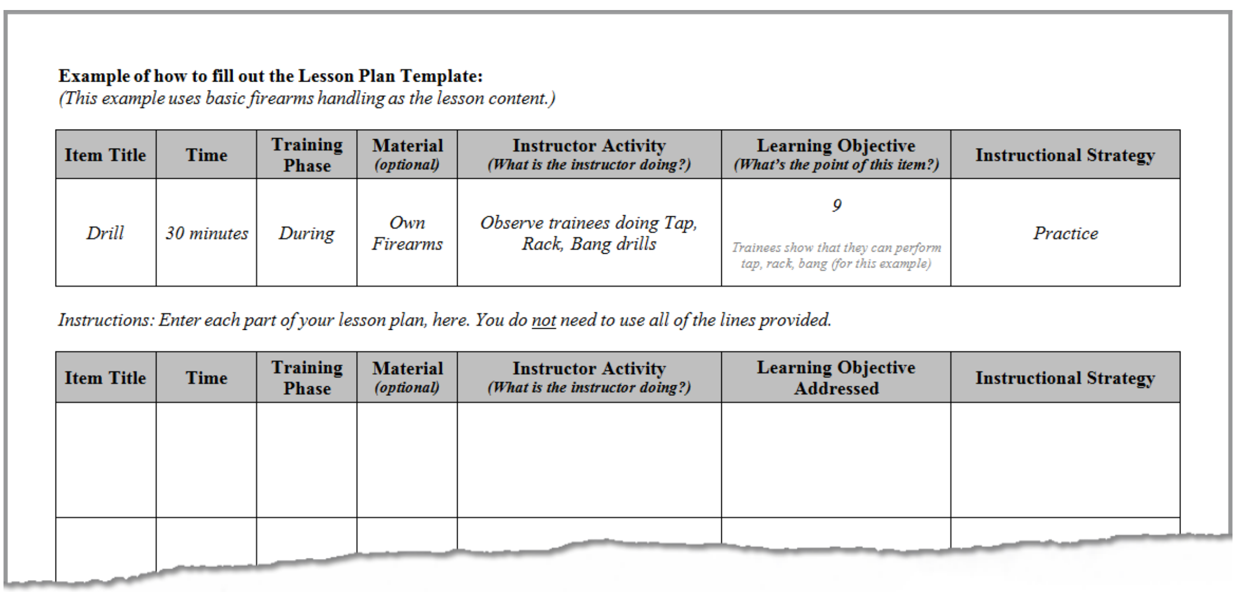
To assess their perception of workload, participants were given the Task Load Index developed by NASA, which is considered a sensitive and reliable measure of workload (Hart & Staveland, 1988). It separates workload into mental demand, physical demand, temporal demand, performance, effort, and frustration. Each of these factors is rated on a seven-point scale.

Instructional Knowledge Posttest

After authoring their lesson plans, participants completed a posttest on instructional knowledge in order to assess whether exposure to the ISS enhanced participants' enduring knowledge of instruction (as compared to the control group). The 14-item multiple choice posttest asked questions regarding instructional strategy use (e.g., *An important purpose of instructional strategies is to…?*) and simulation-based training (e.g., *Which of the following is not a real simulation use?*).

Reactions Survey

Finally, participants completed a 12-item, Likert-style subjective reaction survey that examined their perceptions of the DVTE simulation system (control) or the ISS software (experimental).



Example of how to fill out the Lesson Plan Template:
(This example uses basic firearms handling as the lesson content.)

Item Title	Time	Training Phase	Material *(optional)*	Instructor Activity *(What is the instructor doing?)*	Learning Objective *(What's the point of this item?)*	Instructional Strategy
Drill	*30 minutes*	*During*	*Own Firearms*	*Observe trainees doing Tap, Rack, Bang drills*	*9* *Trainees show that they can perform tap, rack, bang (for this example)*	*Practice*

Instructions: Enter each part of your lesson plan, here. You do not need to use all of the lines provided.

Item Title	Time	Training Phase	Material *(optional)*	Instructor Activity *(What is the instructor doing?)*	Learning Objective Addressed	Instructional Strategy

Figure 5. Lesson Plan Template Apparatus

Results

Prior Instructional Experience

Across all the participants, 60 claimed to be directly responsible for conducting and/or supervising the training of other Marines. Thirty-four of these 60 indicated that they had received some type of training on how to train others. No significant differences existed between the experimental and control groups. Prior instructional experience variables did not explain any additional variance in outcome results.

Call For Fire Pretest

A randomly occurring selection effect was found between the experimental and control groups' pretest scores. A one-way ANOVA indicated that the control group (M = 8.325) performed significantly higher (p < .05) on the pretest than the experimental group (M = 7.1). Consequently, these pretest scores were used as a significant covariate when evaluating lesson plan outcomes.

Lesson Plans Assessment Rubric

Four experimenters graded the participants' lesson plans against a rubric developed *a priori*. Experimenters assessed lesson plan quality in three categories: instructional strategy use, inclusion of supporting training and education events, and learning objective alignment. More specifically, the 53-item rubric primarily included prompts scored as pass-completely, pass-partially, or fail for each of the eight lesson objectives specified in the participant instructions (e.g., *Was a declarative knowledge strategy included with this objective?*). In addition to these items, graders were asked to make overall assessments of each lesson plan (also on a pass/partial-pass/fail scale) and provide their personal, subjective feedback on four 5-point Likert items. Note: We felt that this rubric may be useful to other experimenters, and as such, have provided an abbreviated version in the appendix of this paper.

<u>Rubric Operational Validity</u>. Before the experimental team graded the participants' lesson plans, we asked E-PME subject matter experts to review the lesson plan grading rubric. Four E-PME instructional staff members each graded two sample lesson plans, and then provided subjective feedback about the utility and validity of the rubric. We used their feedback to refine the accuracy and usability of the rubric before commencing grading.

<u>Lesson Plan Inter-Rater Reliability</u>. Four raters independently graded each lesson plan against the rubric. To establish inter-rater reliability, five plans from the control and experimental groups were first graded by each of the raters and then compared. A one-way intra-class reliability analysis indicated ICC (1,3) = .940. The raters then discussed their grading and agreement was reached on several definitions. After all plans were graded, a second analysis indicated a reliability of .955 across the four raters for all 80 participants.

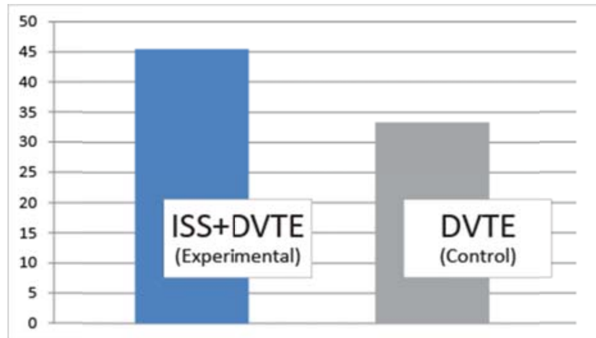


Figure 6. Lesson Plan Outcome Scores

Lesson Plan Outcomes

The lesson plan scores were averaged across the four raters for each participant. An ANCOVA, using pretest Call For Fire scores as a covariate (p < .05), indicated a significant difference between the control and experimental groups $F(3,79) = 2.07$ (see Figure 6). The experimental group (M = 45.39) scored significantly higher on the lesson plan than the control group (M = 33.33). The effect size (d = .54) indicates a medium effect for this result. Or, stated more plainly, the experimental group, using the ISS, performed 36% better than the control group.

Instructional Knowledge Posttest

Posttest instructional knowledge scores were analyzed using a one-way ANOVA, which did not indicate a significant difference between the experimental and control groups. An additional test, using ANCOVA with pretest scores as a non-significant covariate, also did not indicate any significant differences between the experimental and control groups.

Subjective Workload Measure

TLX responses were calculated for all participants in both groups to determine overall perceived workload. A range of 6 to 36 is possible where the higher the total the greater the perceived task load. Although, the TLX includes a weighting scheme, contemporary research suggests that weighted TLX scores provide no additional predictive power compared to non-weighted scores (e.g., Tsang & Wilson, 1997). Consequently, weighted scores were not used in this analysis; only raw scores were analyzed. A one-way ANOVA indicated that no significant differences existed between the overall mean scores of the experimental and control groups. Therefore, both groups considered the task to be equally demanding in regard to workload.

Reactions Survey

Subjective responses were calculated for both groups, and higher scores indicate more positive reactions. A one-way ANOVA indicated that significant differences existed between the experimental group and control group. The control group (M = 43.0) responded with more positive overall reactions than the experimental group (M = 34.5).

Discussion

These data suggest that the ISS software allowed unofficial instructors (i.e., Marine sergeants) to construct more instructionally sound lesson plans. This result was uncovered even though all participants received a lesson plan template that explicitly prompted them to consider key ISD principles (effectively providing a positive intervention to the control cohort, as well as, the experimental group). This outcome suggests that, as we hypothesized, the ISS software can support informal simulation instructors.

However, participants in the control group reported significantly higher positive reactions to the DVTE-only condition. Based upon informal participant comments, we suspect that the presence of DVTE manuals, and comparative lack of ISS user manuals, led to this response. Thus, despite the lack of user manuals, the ISS condition outperformed the control condition, but (perhaps not surprisingly) experimental participants expressed more frustration regarding the process.

In summary, the ISS appears to successfully support informal instructors by helping them more effectively design lesson plans for simulation-based training.

CONCLUSION

The NEW-IT team set out to research, design, test, and evaluate an Instructional Support System to support two broad classes of simulation-based training instructors: officially trained billeted instructors and mostly untrained informal instructors. Throughout the three-year development effort, the utility of each iterative prototype was evaluated. In the end, the final prototype (V2.5 of the anticipated 3.0 versions) demonstrated 79% improvement in operational efficiency (i.e., in scenario setup time) and 98% improvement in operational effectiveness (i.e., instructors' ability to accurately select specific scenarios).

While the operational assessment team assessed the utility of ISS prototypes, the instructional effectiveness team investigated the software's impact on learning outcomes (i.e., instructional execution) and lesson plan quality (i.e., instructional planning).

The first field assessment examined whether the instructional strategies built into the ISS (and validated in the laboratory) would demonstrate significant positive effects on knowledge acquisition in field settings. That experimental investigation found that the ISS led to 26–52% improvements in knowledge outcomes, depending upon the type of knowledge evaluated (see Vogel-Walcutt et al., 2011). The second instructional effectiveness assessment, described in this paper, sought to uncover whether instructors (particularly, informal instructors) could use the ISS to build more instructionally sound training episodes. The results of this investigation suggest that Marine sergeants were able to build significantly more effective lesson plans with the software. In this particular instance, ISS improved lay-instructors' lesson plans by 36% versus the control condition.

The lessons learned in these investigations were used to inform subsequent versions of the ISS. The final prototype system, version 3.0, was delivered to the Office of Naval Research in Summer 2011. The ISS is now considered Government Off-The-Shelf (GOTS) software.

ACKNOWLEDGEMENTS

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Department of Defense or Office of Naval Research. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon. This work was supported by the Office of Naval Research Grant N0001408C0186, Next-generation Expeditionary Warfare Intelligent Training.

REFERENCES

Amos, D. (2010). *Eclipse of the Sunnis.* Washington: Public Affairs.

Becker, W. & Schatz, S. (2010). Higher order cognitive skills training in support of contemporary military operations. In proceedings of the human factors and ergonomics society annual meeting (2228-2232). San Francisco, CA. 10.1518/107118110794001090.

Combat Instructor School Overview (n.d.). Camp Geiger, NC.

Fowlkes, J., Stagl, K., Schatz, S., Taylor, A., & Lum, H.

(2011). Effects of instructional strategy guidance on instructor knowledge and quality of instruction (under review).

Hart, S.G., & Staveland, .E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology, 52*, 139-183. 10.1016/S0166-4115(08)62386-9.

Krulak, C. C. (1999). The Strategic Corporal: Leadership in the Three-Block War. *Marines Magazine, January*.

Loftin, R. B., Kenney, P., & Mastaglio, T. (2004). *Outstanding research is-sues in intelligent. tutoring systems.* (CDRL A003). Portsmouth, VA: U.S. Army RDECOM.

Morrison, J. E., & Fletcher, J. D. (2002). Cognitive Readiness. Institute for Defense Analyses. Alexandria, Virginia

Oser, R. L. Gualtieri, J. W. Cannon-Bowers, J. A., & Salas, E. (1999). Training team problem solving skills: an event-based approach. *Computers in Human Behavior, 15*(3-4), 441–462.

Ross, K. G., Phillips, J. K., Klein, G., & Cohn, J. (2005). Creating Expertise: A Framework to Guide Simulation-Based Training. In *Proceedings of I/ITSEC*, Orlando, FL: NTSA.

Salas, E., Rhodenizer, L., & Bowers, C. A. (2000). The design and delivery of crew resource management training: Exploiting available resources. *Human Factors, 42*(3) 490–511.

Schatz, S., Lackey, S., Stanney, K., & Schaffer, R. (2011). Next-generation expeditionary warfare intelligent training (NEW-IT). (In press).

Tsang, P.S., & Wilson, G.F. (1997). Mental Workload. *Handbook of Human Factors and Ergonomics* .(Chap. 13). New York: John Wiley.

Vogel-Walcutt, J.J., Carper, T.M., Bower, C., & Nicholson, D. (2010). Increasing efficiency in military learning: Theoretical considerations and practical applications. *Military Psychology*, *22*(3), 311-339.

Vogel-Walcutt, J., Marshall, R., Fowlkes, J., Schatz, S., Dolletski-Lazar, R., & Nicholson, D. (In Press, 2011). Effects of instructional strategies within an instructional support system to improve knowledge acquisition and application. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* Las Vegas, NV.

APPENDIX

Training Plan Grading Rubric (Abbreviated)

Instructional Strategy Use

Total number of **different** instructional strategies listed? ______________________________

Total number of instructional **events** listed (i.e., total rows)? ______________________________

Inclusion of Supporting T&E Events

Each item below should be scored as either *yes* or *no*. However, if in rare cases "partial credit" should be awarded, mark *somewhat*.

Inclusion of "**pre-training**" (aka "preparation" or "pre-brief") .... Yes – No – (Somewhat)
Inclusion of "**during-training**" (aka "practical application/exercise") .... Yes – No – (Somewhat)
Inclusion of "**post-training**" (aka "reflection" or "AAR") .... Yes – No – (Somewhat)
Inclusion of **simulation-based training** (aka a "scenario" or "event-based") .... Yes – No – (Somewhat)
Inclusion of an **assessment** (e.g., test, simulation-based exercise) .... Yes – No – (Somewhat)

Learning Objective Alignment

Each item below should be scored as either *yes* or *no*. However, if in rare cases "partial credit" should be awarded, mark *somewhat*.

Learning Objective 1: Definition of CFF (Declarative)

Was this **learning objective included** in the lesson plan? .... Yes – No – (Somewhat)
Was **any instructional strategy** associated with this objective? .... Yes – No – (Somewhat)
Was a **declarative knowledge** strategy included with this objective? .... Yes – No – (Somewhat)
Was **novice-level strategy** associated with this objective? .... Yes – No – (Somewhat)
Did the **instructor activity match** the instructional strategy? .... Yes – No – (Somewhat)
Did the listed strategy correspond the listed **training phase**? .... Yes – No – (Somewhat)
Was this objective associated with the **pre-training phase**? .... Yes – No – (Somewhat)

SME Subjective Assessment

Each item below on a scale of 1-5, where 1 is the worst and 5 the best.
1 = Very poor 2 = Poor 3 = Fair 4 = Good 5 = Excellent

Rate the **completeness** of the overall lesson plan? .... 1 – 2 – 3 – 4 – 5
Rate the **correctness** of the overall lesson plan? .... 1 – 2 – 3 – 4 – 5
Rate your personal interest in **using this** exact lesson plan? .... 1 – 2 – 3 – 4 – 5
How would you rate the lesson, **overall**? .... 1 – 2 – 3 – 4 – 5