Evaluating Fidelity for Tactical Training within the Live-Virtual Constructive Environment

Barbara Buck, Sarah Davies,
Tim Stumpf, Rob Lechner
The Boeing Company
St. Louis, MO
barbara.j.buck@boeing.com
sarah.davies4@boeing.com
tim.stumpf@boeing.com, rob.lechner@boeing.com

Winston Bennett, Jr.
Air Force Research Laboratory
Wright-Patterson AFB, OH
winston.bennett@us.af.mil

ABSTRACT

Advances in Live Virtual Constructive (LVC) technology will soon make it possible for aircrew in the Combined Air Forces to utilize LVC capabilities on a daily basis. LVC is enabled on a live platform by integrating various virtual sensor models into the weapon system that replicate the functions of live sensors onboard the weapon system. These sensor models can be implemented at various levels of fidelity, from simplistic range bin detection models to physics-based, environment-enabled models. This paper describes a research experiment conducted in July 2020 to measure aircrew sensitivities of various levels of sensor model fidelity used on a platform within an LVC framework. This effort builds upon existing LVC capabilities developed for the F-15E platform under the Secure LVC Advanced Training Environment (SLATE) program (Lechner & Huether, 2008, Lechner & Wokurka, 2010, Lechner & Schwering, 2012, Call & Lechner, 2018). In this study, aircrew were presented targets of varying levels of fidelity for both an electromagnetic sensor (AESA Radar) and an electro-optical sensor (Sniper Targeting Pod) utilizing the sensor hardware on an F-15E avionics bench. The 18 participants were current and former pilots and Weapon System Officers (WSO). The impacts of sensor fidelity were evaluated using a combination of objective performance metrics and subjective aircrew ratings for target detection and identification tasks. Subjective responses of participants indicated favorable ratings regarding the LVC capability for emulating real-world conditions to support training for all tasks. Participants reported that functionality was realistic, workload was manageable, and the systems were acceptable to meet training needs. Objective performance measures generally supported the more true-to-life targets for tactical tasks, however other instances supported a mix of sensor fidelities. This paper details all experimental results, gives recommendations for specific implementations within the LVC environment, and ideas for future research directions are discussed.

ABOUT THE AUTHORS

Dr. Barbara Buck is a Learning Scientist in the Boeing Global Services Engineering Test and Technology Group. In this role, she explores technological advances to create innovative training solutions and conducts training effectiveness research as training is deployed. Dr. Buck has over 30 years' experience at Boeing, with a focus on adaptive learning, virtual and augmented reality, data analytics, and applications for improved training effectiveness, performance, workload assessments and usability evaluations. She has a M.A. and Ph.D. in Cognitive & Engineering Psychology from the University of Illinois.

Dr. Sarah Davies investigates emerging technology and training solutions as a learning scientist for the Boeing Research & Technology organization in St. Louis, MO. She earned a M.S. in Instructional Design and Educational Technology and a Ph.D. in Learning and Cognition from the University of Utah. Dr. Davies has 18 years of experience working in education and training across a variety of fields; her areas of expertise include human computer interaction, comprehension, digital learning, and accessibility.

Tim Stumpf is a Senior Engineer on the Advanced Learning team for the Boeing Research & Technology (BR&T) organization in St. Louis, MO. Tim is the lead systems engineer for the BR&T Secure Live Virtual Constructive Advanced Training Environment (SLATE) program. He has more than 35 years' experience at the Boeing Company and has received 3 patents in the area of aircrew training technologies. He received a Bachelor of Science degree in Business from Columbia College.

Rob Lechner is the Chief Engineer of the Advanced Learning team for the Boeing Research & Technology organization in St. Louis, MO and an Associate Technical Fellow in the field of Live, Virtual, Constructive technology. He is also the Program Manager for the BR&T Secure Live Virtual Constructive Advanced Training Environment (SLATE) program. He has more than 30 years' experience in aircrew and distributed mission training advancements and has received 11 patents in the area of aircrew training technologies. His education includes a Master of Science degree in Engineering Management from the Missouri University of Science & Technology; Bachelor of Science degree in Electrical Engineering and Communication Sciences from the University of Central Florida.

Dr. Winston "Wink" Bennett is a Senior Principal Research Psychologist and Readiness Product Line Lead for the Warfighter Interactions and Readiness Division, Airman Systems Directorate, 711th Human Performance Wing, Air Force Research Laboratory. Wink is a recognized leader in education, training, competency definition and assessment and performance measurement research. He has been involved in a number of Joint and Multinational research collaborations and continues to support science and practice collaborations around the world.

Evaluating Fidelity for Tactical Training within the Live-Virtual Constructive Environment

Barbara Buck, Sarah Davies,
Tim Stumpf, Rob Lechner
The Boeing Company
St. Louis, MO
barbara.j.buck@boeing.com
sarah.davies4@boeing.com
tim.stumpf@boeing.com, rob.lechner@boeing.com

Winston Bennett, Jr.
Air Force Research Laboratory
Wright-Patterson AFB, OH
winston.bennett@us.af.mil

INTRODUCTION

With the advent of synthetic environment injects into live aircraft soon to be realized within tactical air platforms, the technical capability will exist for aircrew to train with a Live, Virtual, Constructive (LVC) capability on a recurring basis. LVC is enabled on a live platform by the addition of various sensor simulation models into the weapon system. These models replicate the functions of live sensors onboard the weapon system, providing track information from the synthetic environment that is merged with the live environment. These sensor models can be designed at various levels of fidelity, from simplistic range bin detection models to physics-based, environment-enabled models. There is a direct proportional relationship between level of fidelity increase and impacts to central processing unit (CPU) and graphics processing unit (GPU) processing power necessary to execute the sensor algorithms. This in turn, impacts the requirements of the computational power necessary for an LVC Program of Record. A high-fidelity, ray-tracing physics-based model may provide a result exacting a live sensor, but the processing power required to operate it becomes overwhelming. The need exists to determine how much fidelity is sufficient to meet the training objectives.

The study Fidelity Integration within Tactical Training of LVC (FITTL), sponsored by the Air Force Research Laboratory (AFRL), was conducted to provide an initial look into the impact of a subset of aircraft sensor models on performance and perceptions during training within an LVC environment. An experimental bench was assembled in the F-15E Electronic Systems Integration Laboratory (ESIL) in St. Louis, MO utilizing actual aircraft hardware and avionics. The experimental bench provided the ability to utilize a combination of actual sensor hardware and various levels of fidelity of sensor models allowing the experimental protocol to vary the level of fidelity presented to aircrew. This paper details the experiment bench, the experimental protocol, and the study results from the experiment.

OVERVIEW OF LVC CAPABILITY/SYMBOLOGY DEVELOPMENT

LVC Capability Overview

Boeing utilized the F-15E Operational Flight Program (OFP) developed under the AFRL SLATE Program for this study. The SLATE OFP supports LVC operations for both an Active Electronically Scanned Antenna (AESA) radar model and a Sniper Targeting Pod (TPOD) model. Both the AESA model and the Sniper model are contained within the LVC Processor Module (LVCPM), an internal component of the SLATE modified P5 Tactical Combat Training System (TCTS) pod, see Figure 1.

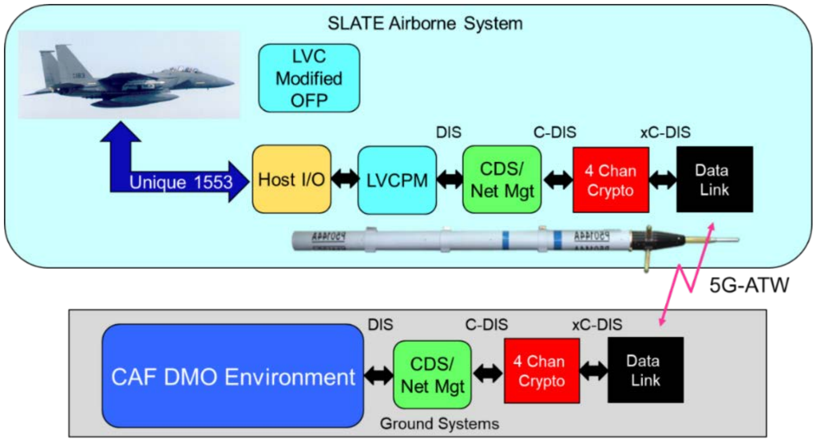

Figure 1. Project SLATE LVC Airborne System

[F-15 photo credited to Boeing, P5 Pod credited to Leonardo DRS]

The test setup in the F-15 ESIL is shown in Figure 2. The color key defines the existing AESA bench equipment, the SLATE equipment, and the new equipment added for the study. The LVC Processor Module (LVCPM) was emulated on a desktop computer for this study and included the LVC Operational Executive Program (OEP) software, Big Tac, the AESA model, the TPOD model, and the Host Input/Output (I/O) emulator. The Advanced Display Core Processor (ADCP) executed the F-15E SLATE modified OFP. Both the AESA RADAR and the TPOD avionics subsystems were utilized as inputs to the experiment. The real AESA and TPOD (supports live tracks) are used in conjunction with an AESA model and a TPOD model (supports virtual/constructive tracks) to produce an LVC environment in the cockpit. The real TPOD is situated in a window in the ESIL facility such that it is facing Lambert airfield in St. Louis. Specific for the TPOD for both Air-to-Air and Air-to-Ground modes, video images of synthetic infrared targets were generated on a utility computer using GenesisIG™ software. The simulated targets were combined with real TPOD video using the EZwindow™ device to produce an augmented reality display in the cockpit. The resultant raster video is a combination of real infrared imagery from the TPOD with virtual targets overlaid on the real-world imagery. While this approach has the limitation of not providing appropriate altitude offsets and slant angles for Air-to-Ground, we were able to locate targets at appropriate distances to minimize this affect. For Air-to-Air, targets were positioned is the sky such that the TPOD video portray only sky imagery.

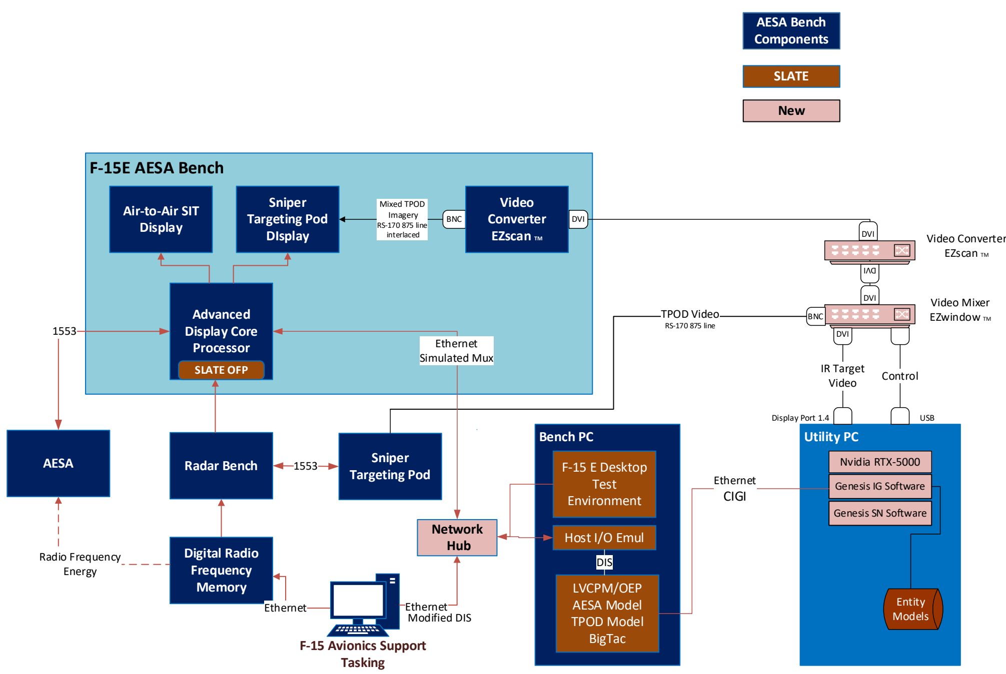

Figure 2. Test System Diagram

ESIL Equipment

Boeing developed and maintains an AESA test bench that resides in the F-15 ESIL. The test station houses the core avionics processors, displays, and cockpit controls of the F-15E and is comprised of both real flight hardware, such as the ADCP and cockpit stick and throttle. The station contains a combination of real and simulated displays and software-defined sub-system models. The union of real flight hardware and software-simulated models provides a cost-effective means to exercise system functions and integrate new features in a controlled environment. The TPOD is an electro optic sensor containing both a Forward Looking Infrared detector and a Charge Coupled Device Television detector to generate raster video for display in the cockpit. The TPOD also includes image trackers for Line of Sight stabilization, a laser transmitter/receiver for ranging and designating targets for laser guided bomb deliveries, a night vision goggle compatible laser marker for designating and illuminating targets, and a laser spot tracker.

LVCPM

Boeing designed and manufactured the LVCPM to provide additional computer processing capability outside of the aircraft mission computer to support LVC training missions. In addition, we designed, built and tested the LVCPM to fit within the form factor of the P5 TCTS Pod envelope as well as an option for an internal carriage configuration inside the aircraft. The LVCPM provides a flexible solution that enables high fidelity training content delivery in support of embedded training and LVC training missions. The LVCPM was designed to survive the rigors of a military flight environment and is flight qualified as a component of the SLATE configuration of the P5 TCTS pod. The LVCPM contains up to six (6) ruggedized Quad Core Atom Processor modules (single board computers). To simplify testing equipment and testing architecture, the LVCPM device was not used during this study however, the LVCPM software was hosted on a standard Dell desktop computer emulating the performance of the LVCPM. The LVCPM software is an open architecture design that provides the synthetic environment for a live aircraft. An executive layer moderates the execution and communications of applications and plug-ins that reside across the internal processors within the LVCPM. The LVCPM communicates with the rest of the SLATE components over the IEEE 1278.1 Distributed Interactive Simulation (DIS) protocol.

Software Elements

Two new software components were developed for this study: (1) the DIS to Common Image Generator Interface (CIGI) converter and (2) the TPOD CIGI Controller. The DIS to CIGI converter is a library that converts general DIS protocol messages and translates them to CIGI messages to inform the Image Generator of the visual environment to draw. The computer generated threats and the "live" own ship or F-15E bench are the sources for that visual environment data. The TPOD CIGI controller receives the F-15E TPOD model current state and converts the state data to CIGI messages for proper reflection in the visual environment.

Constructive Force Generator

Big Tac is a flexible, high-fidelity threat environment capable of presenting a combination of air threats and ground based air defense threats to enhance immersion of trainees into a synthetic combat environment. It is the standard threat environment for USAF F-15 and F-16 Distributed Mission Operation training systems. It is designed to meet single-ship, multi-ship, local, and long haul networking requirements between dissimilar aircraft types in support of joint, combined, and coalition training. Big Tac provides a rich set of stimuli for weapon, avionics and visual systems used by the trainees. It simulates a mix of interactive entities, accurately modeling their physical, behavioral, electronic, and countermeasure capabilities. Big Tac is a commercial product of the Boeing Company. Big Tac will be delivered with a graphical scenario generation tool for developing training scenarios and a threat library maintenance tool to aid management of customer specific threat performance data and algorithms.

Symbology Development

For this study, we used the SLATE-developed AESA radar model as the baseline model and developed two additional levels of fidelity to provide radar functionality closer to the behavior of the real radar system. Test subjects addressed constructive threats presented on the AESA radar display. In addition, for the electro-optical target presentation, we developed three levels of visual model fidelity target representation on the TPOD display, in both the air-to-air and the air-to-ground modes. The TPOD representations included a reduced or simplified polygonal 3D model, a high polygon count polygonal 3D model, and a high polygon count polygonal 3D model with infrared (IR) effects. A scene generator synchronized the 3D models with the external environment and displayed them together on the cockpit TPOD display.

Multiple mission scenarios were developed with the intention of exposing the test subjects to the levels of fidelity changes that were made to the AESA radar model and the TPOD model. Mission scenarios were developed using Big Tac for constructive force generation. Test trials were developed using the Big Tac mission scenarios combined with the different levels of fidelity models. Test trials were grouped together into three (3) primary tasks. Test subject participants completed the Air-to-Air Radar Detection and Identification task, the TPOD Air-to-Ground Identification task, and the TPOD Air-to-Air Detection and Identification task during the course of the testing.

Visual Models

Image Generation (IG) software from Diamond Visionics, along with a Sensor Module also from Diamond Visionics for physics-based simulation of IR effects, was employed in this study in order to render simulated 3D targets for injection into the raster video from the TPOD. The Target Image Generator is configured to receive CIGI commands as input from the LVCPM TPOD CIGI controller and render the simulated targets correlated to the real world coordinates in real-time. During FITTL testing, the skybox and terrain are masked out of the rendered scene so that the target appears with a black background. This allows for the simulated targets in the scene to be combined with the TPOD video using Luma keying.

Video blending hardware was used in this study to combine simulated targets, generated by the IG software, with video from the TPOD. A Westar EZwindow™ video blender and a Westar EZscan™ scan converter was employed to accomplish video blending.

A set of 3D OpenFlight models were selected from the Boeing Common Data Set (BCDS) for this study. These selected models were used to create a corresponding set of low, medium, and high fidelity models. (see Table 1). The lowest level of fidelity for each model was chosen such that a low end computer could render the images. The medium fidelity model included the highest polygon count of the model. The high detail model was the same polygon count of the medium fidelity model with the addition of texture pattern sensor attribution such that it could be rendered by the GenesisSN™ software. Examples of visual model level of fidelity are shown in Figure 3.

Table 1. Level of Fidelity of Visual Models

Level of Fidelity	Model Set	IR Simulated
Level 1: High Visual Fidelity	High Polygon Count with Sensor Attribution	Yes
Level 2: Medium Visual Fidelity	High Polygon Count	No
Level 3: Low Visual Fidelity	Low Polygon Count	No

Figure 3. Examples of Visual Model Level of Fidelity

Radar Models

Updates were made to the AESA radar model to increase the fidelity of the models performance. Changes were made to 1) the time needed to establish a track and 2) the chances of dropping a track after establishment. To accomplish these changes, modifications were made to increase the workload of the radar antenna, the number of simultaneous tasks the radar system is performing and increasing the duration of those tasks. Additionally, the base detection range performance of the model was decreased. The updates to the radar model were developed, reviewed and validated using inputs from pilot subject matter experts. Three selectable model degradation levels were developed that enable model performance to be varied during test missions. The FITTL radar model degradation value can be set to none (baseline SLATE model, perfect detection and no track dropping), moderate (some atmospheric clutter, moderate time to establish tracks with some track dropping), or high (additional atmospheric clutter, longest time to detect tracks and high rate of track dropping).

DESCRIPTION OF EXPERIMENTAL METHODS

Research Questions

This study examined the impact of radar or visual fidelity manipulations on human performance and perceptions of the training experience within an LVC environment. Specifically, it sought to answer these questions:

1. How does the level of radar degradation (none, moderate, high) or visual fidelity (high, medium, low) impact human performance?
    a. Entity detection
    b. Entity identification
    c. Time on task (where applicable)
2. How does the level of radar fidelity impact self-reported perceptions of the training experience?
    a. Confidence in performance accuracy
    b. Sense of time pressure
    c. Acceptability of the training system
    d. Realism of the training system
    e. Workload

Participants

The study was designed for USAF participants of varied experience levels, but due to COVID restrictions all recruiting was limited to Boeing employees. Eighteen (18) current or former operational pilots and weapon system officers (WSO) were recruited from within the Boeing population. These participants were highly experienced with a variety of tactical aircraft platforms, including derivatives for the F-15, F/A-18, F-16 and F-22. Average tactical hours for this group of participants was just over 2600.

Protocol

As noted above, the experiment was carried out in an AESA test bench that resides in the F-15 ESIL and is comprised of both real flight hardware, such as the ADCP and cockpit stick and throttle. The station contains a combination of real and simulated displays and software-defined sub-system models. The union of real flight hardware and software-simulated models provides a cost-effective means to exercise system functions and integrate new features in a controlled environment. The experiment consisted of a pre-experimental phase and three experimental tasks. In the pre-experimental phase, participants completed informed consent and a brief demographics questionnaire. During the experimental phase, participants performed three different tasks: 1) Air-to-Air Radar Detection and Identification, 2) Sniper Targeting Pod Air-to-Ground Identification, and 3) Sniper Targeting Pod Air-to-Air Detection and Identification. Performance measures, subjective ratings, and open-ended feedback were collected during each task. Each participant was tested individually and took approximately four (4) hours to complete the testing.

Subjective Ratings

Across all tasks, subjective ratings were collected from participants following completion of each individual trial. Participants were asked to rate their agreement with a set of four statements on a labeled five-point Likert scale, from Strongly Disagree (1) to Strongly Agree (5). These statements gauged participants' 1) confidence in their performance accuracy, 2) sense of time pressure during the trial, 3) perceived acceptability of the training system, and 4) perceived realism of the training system. An additional Likert-scale item asked participants to provide a self-assessment of their workload during the trial using another labeled five-point scale, ranging from Under-utilized (1) to Excessive (5).

Air-to-Air Radar Detection and Identification Task

The first task involved detecting targets on the situation display (see Figure 4), and identifying targets as “friendly” or “hostile” to differentiate between noise, clutter, and actual threats. Test scenarios for the air-to-air radar task were created using the Big Tac constructive force generation tool. For the air-to-air radar trials, we used three (3) levels of scenario complexity (low, medium, and high) and three (3) levels of radar model degradation: none, moderate and high. Each of these levels was completely crossed, resulting in nine (9) test conditions. Each participant completed two (2) different trials at each difficulty by radar model combination, resulting in eighteen (18) total trials per participant. The order of trials was completely randomized for each participant.

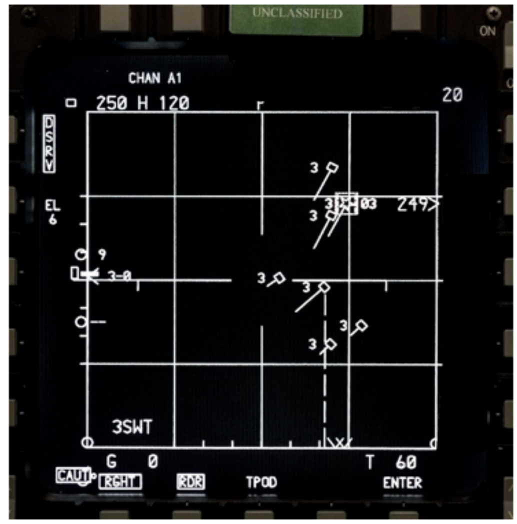

Figure 4. Tactical Entities on the Radar Display

Following each trial, participants rated how realistic and acceptable the radar models were for the radar detection and identification tasks (subjective assessments), and rated subjective workload (modified from Kirwan, et. al, 1997).

Sniper Targeting Pod Air-to-Ground Identification Task

For the second task, participants searched for and identified constructive imagery presented against the real world background using an air-to-ground targeting pod. The computer-generated entities were placed in pre-determined locations within the limited FOV of a stationary TPOD (see Figure 5). Nine (9) different entities comprised the air-to-ground trials: four (4) friendlies (car, jeep, tractor trailer, Bongo without gun) and five (5) threats (Bongo with gun, Humvee, SA-11, scud launcher, ZSU 23). For each trial, a subset of five (5) of the possible targets was presented. Each participant completed fifteen (15) trials; five (5) of 15 possible Big Tac scenarios for each of three image fidelity levels. Participants were given time to familiarize themselves with the different entity types prior to the identification task. Once all five (5) entities were located and identified for a given trial, the participant provided subjective ratings similar to those in the air-to-air radar task. The attending experimenter collected and recorded measures of response time and identification accuracy for each trial.

Figure 5. Examples of Ground Entities on TPOD display

Sniper Targeting Pod Air-to-Air Detection and Identification Task

As with the air-to-ground task, the air-to-air task involved searching for and identifying constructive entities presented against the real world background using an air-to-ground targeting pod. In order to locate the air entities with the TPOD, participants first locked onto a radar track, and then commanded the TPOD to the radar track. Participants then identified the different entities. The threats used in this task were Mig-29 and SU27, friendlies were the F-15 and F-16. The total number of entities per air trial ranged from two (2) to eight (8), depending on scenario difficulty. Participants completed nine (9) total air-to-air TPOD trials, one (1) of each difficulty level for each of the three (3) imagery fidelity levels. Time on task (capped at a maximum time of three (3) minutes per trial), identification accuracy, and subjective ratings were collected after each trial.

SUMMARY OF RESULTS

Air-to-Air Radar Outcomes

Performance measures for the Air-to-Air Radar Detection and Identification Task included the percent of total entities detected as well as the percent of entities correctly identified as friend or foe. Additionally, subjective ratings of radar degradation provided additional insight. Results indicated a significant main effect of radar degradation on detection (F = 37.0, p < .001); detection rates were highest with no radar degradation and lowest with high degradation. There was also a significant main effect of scenario difficulty (F = 26.0, p < .001), showing the expected pattern of higher detection in low and medium difficulty scenarios and lower detection in higher difficulty scenarios. The interaction of radar degradation and scenario difficulty on entity detection approached significance (F = 2.0, p = .09). Whereas detection was excellent for low and medium difficulty tasks with no or moderate radar degradation, the pattern for high radar degradation deviated from this overall trend, achieving lower detection rates in low difficulty scenarios than in medium difficulty scenarios. Results for identification accuracy followed a similar pattern (See Figure 6). There was a significant main effect of radar degradation (F = 51.3, p < .001); again, accuracy was highest with no radar degradation and lowest with high degradation. There was a significant main effect of scenario difficulty (F = 19.96, p < .001), with the highest accuracy achieved in medium difficulty scenarios. There was a significant interaction of radar degradation and scenario difficulty (F = 3.2, p < .05). While increasing scenario difficulty with no radar degradation showed a decreasing trend in identification accuracy, identification accuracy for both low and high difficulty scenarios was lower than for medium difficulty scenarios with moderate and high radar degradation.

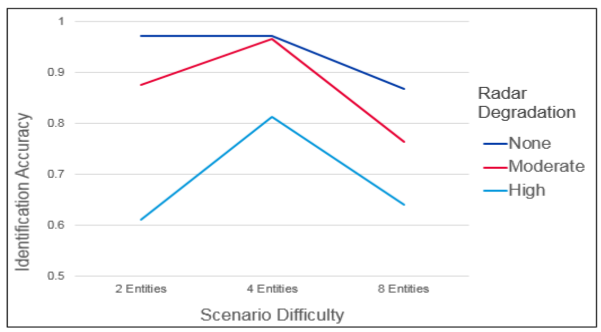

Figure 6. Identification Accuracy as a Function of Scenario Difficulty and Radar Degradation

A Chi-square test of independence was conducted to investigate the relationship between radar degradation and participants' reported confidence in their task performance. The test results indicated a significant effect of radar degradation on participants' responses (X^2 = 13.8, p < .01), with significant contributions to the Chi-square value from the Agree and Disagree categories of the high degradation condition. Participants less often agreed and more often disagreed that they felt confident in their performance following trials with high radar degradation. Chi-square tests conducted to examine the relationships between radar degradation and participants' subjective ratings showed no significant impacts on reports of time pressure, radar system realism, task workload, or acceptability of the radar system for training.

Sniper Targeting Pod Air-to-Ground Outcomes

Performance measures for the TPOD air-to-ground task included task time and the percent of entities correctly identified. Additionally, subjective ratings examined the participants' evaluation of acceptability for varying levels of target fidelity. Results showed that the effect of target fidelity on task time approached significance ($F = 2.6$, $p = .08$); time to completion was shortest with medium fidelity (Mmedium = 37 seconds) and longest with high fidelity (Mhigh = 45 seconds). This effect is illustrated in Figure 7. The measure of identification accuracy showed a ceiling effect with all conditions achieving a mean accuracy rate of 97% or above; there was no significant effect of target fidelity.

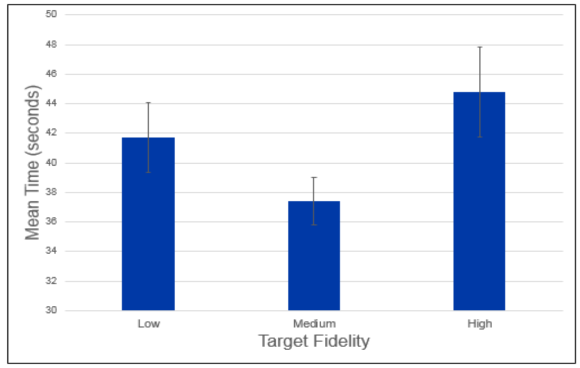

Figure 7. Air to Ground TPOD Identification Time as a Function of Target Fidelity

A Chi-square test of independence investigating the relationship between target fidelity and participants' perception of acceptability of the TPOD entities for training was not significant; however, a Chi-square test comparing only the medium and high fidelity conditions did indicate an effect ($X^2 = 7.8$, $p = .05$). Participants less often strongly agreed and more often strongly disagreed that the TPOD entities were acceptable following trials with high fidelity than with medium fidelity. Chi-square tests conducted to examine the relationships between target fidelity and participants' subjective ratings showed no significant impacts on reports of the realism of TPOD functionality, confidence in performance accuracy, time pressure, or task workload.

Sniper Targeting Pod Air-to-Air Outcomes

Performance measures for the TPOD air-to-air task included the percentages of entities detected and correctly identified, as well as task time. Two-way ANOVAs were conducted using the three (3) levels of target fidelity (low, medium, high) and the three levels of scenario difficulty (low, medium, high) to examine participants' performance. Additionally, a series of chi-square tests of independence were performed to examine the relationship between target fidelity and participants' subjective ratings. Results indicated that the effect of target fidelity on detection approached significance ($F = 2.92$, $p = .057$); detection rates were highest with medium fidelity. There was a significant main effect of scenario difficulty ($F = 9.66$, $p < .001$), showing the expected pattern of higher detection in low and medium difficulty scenarios and lower detection in higher difficulty scenarios. The interaction of radar degradation and scenario difficulty on entity detection was not significant ($F < 1$). These findings are illustrated in Figure 8. Results for task time followed a similar pattern: the effect of target fidelity on time to completion approached significance ($F = 2.43$, $p < .05$); task time was shortest with medium fidelity. There was a significant main effect of scenario difficulty ($F = 40.7$, $p < .001$), showing the expected pattern of shortest task times in low difficulty scenarios and longest task times in high difficulty scenarios. The interaction of target fidelity and scenario difficulty was not significant ($F < 1$). There was a significant effect of scenario difficulty on identification accuracy ($F = 4.63$, $p < .05$), with higher accuracy in low difficulty scenarios and lower accuracy in medium and high difficulty scenarios. The effect of target fidelity on identification accuracy was not significant, and there was no significant interaction of target fidelity and scenario difficulty ($F < 1$).

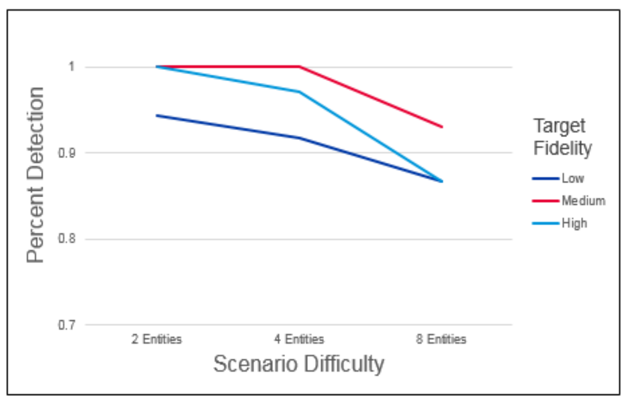

Figure 8. Detection Accuracy as a Function of Scenario Difficulty and Target Fidelity

A Chi-square test of independence was conducted to investigate the relationship between target fidelity and participants' reported confidence in their task performance. The test results indicated a significant effect of target fidelity on participants' responses ($X^2 = 10.8$, $p < .05$), with participants more often agreeing that they felt confident in their performance following trials with medium target fidelity. Chi-square tests conducted to examine the relationships between target fidelity and participants' subjective ratings showed no significant impacts on reports of realism of TPOD functionality, task workload, time pressure, or acceptability of the TPOD entities for training.

Participant Feedback on Tasks

Following completion of all experimental trials, participants provided feedback in response to the question "What did you like or find useful?" An informal thematic analysis of those responses is summarized in Table 2. Overall, we found that a majority of participants commented on the realism of radar and entity behavior, as well as the visual realism of the targeting pod images. Many others referred to the training exercises as a good opportunity for skill development. In response to the question regarding how these exercises could be improved, half described opportunities to increase task difficulty for the radar task and others suggested that ranges needed some adjustment, meaning that entities were detected too far out to be realistic. This was particularly true in the no degradation condition. For the targeting pod tasks, participants thought that the task could be improved with additional target variation and better behavior of the ground images (static in our task). A majority of participants wanted to see improvements to the visual fidelity realism of ground images, and that fidelity of the air images could be degraded further at the far ranges.

Table 2. Themes and Descriptions of Participant Comments by Task

Feedback Type	Theme	Air-to-Air Radar Task	Sniper Targeting Pod Air-to-Ground Task	Sniper Targeting Pod Air-to-Air Task
Strengths	Realism	61% of participants found the radar and entity behavior to be realistic.	50% of participants appreciated elements of authenticity.	33% of participants commented on entity fidelity and visual realism.
	Value of Training	39% of participants described the exercises as a good training exercise for skill development.	39% of participants commented on the value of LVC training.	39% of participants indicated this was a valuable training exercise.
Opportunities	Difficulty	50% of participants identified opportunities for increasing task difficulty.		
	Task Features (Range)	44% suggested that ranges should be closer.	39% suggested that the training task could be improved.	50% of participants suggested a need for more target variation.
	Visual Fidelity & Realism		67% of participants identified opportunities for improving visual fidelity.	28% suggested that the level of fidelity should be adjusted at far range.

CONCLUSION AND NEXT STEPS

Across tasks, participants responded favorably about the LVC capability for emulating real-world conditions to support training. Generally, participants reported that functionality was realistic, workload was manageable, and the systems were acceptable to meet training needs. We used objective performance metrics and subjective ratings at the end of each trial to draw more specific conclusions about model and entity fidelity. Open-ended comments provided additional insights; however, because they were collected only at the conclusion of each task, they do not speak to specific conditions of radar or model fidelity.

Air-to-Air Radar Tasks

Participants responded favorably to all radar models, agreeing that they were realistic and acceptable for training. However, we found that the high radar degradation condition increased task difficulty and more successfully emulated the behavior of real radar. We recommend using the high degradation condition for training applications. Training scenarios for LVC air-to-air radar tasking could be improved by ensuring that radar behavior is appropriate for range and entity behaviors. This could be accomplished by ensuring that dropped tracks occur more frequently at farther ranges and at specific orientations, rather than randomly. Tracks for entities closer in and at head-on orientations should be more stable.

Sniper Pod Air-to-Ground Targeting Tasks

Low-altitude task constraints likely contributed to the ceiling effect on entity identification accuracy, limiting the conclusions that we were able to draw. Based on timing data and subjective feedback, there is evidence that the high fidelity model increased task difficulty and more successfully emulated real-world visualization. This was particularly true when compared to the medium fidelity models. The addition of infrared modeling to the high-fidelity entities helped them blend more with the other real-world objects. Recommendations for future implementations would be to address the instability of entities due to lags in the system, to improve the contrast blending of constructive entities, and to provide the white-hot/black-hot capability available in current targeting pod systems. We expect that the realism factor of ground constructive entities will improve when tasks are performed at more realistic altitudes.

Sniper Pod Air-to-Air Targeting Tasks

Both the low fidelity and high fidelity models produce comparable increases in task difficulty, emulating real-world challenges of detection and identification. This suggests that one potential improvement would be to create a hybrid model that combines lower fidelity models at the farther ranges and higher fidelity infrared models at closer ranges. Furthermore, this task could be improved in future implementations with the addition of the white-hot/black-hot capability available in current targeting pod systems, and by incorporating more diversity in the aspect angles of constructive entities within the scenarios.

Target Image Generator

As the experimental results show, a full-fidelity IR target generator may not increase task complexity to the point of being a requirement. However, a medium fidelity Target Image Generator (TIG) would be beneficial. To that end, it is recommended to pursue the installation of TIG software within the SLATE architecture. The most advantageous system to run the TIG application would be the LVCPM.

Next Steps

There is beneficial research that can be performed to increase the knowledge base of embedded training while minimizing the impact to the host platform. From the results of the study, we recommended several courses of action be taken to 1) improve the realism of radar model performance and 2) continue the study of introducing visual targets in the cockpit of an LVC enabled aircraft.

1. Perform a study/analysis of the entire system (aircraft system, bench system, SLATE system, Image Generator, and/or network) to determine the cause of the transport delay that appeared in the experiment

bench and correct/minimize the deficiency if possible.

2. Perform a study to address Target Pod simulation fidelity. It is recommended that a medium fidelity TIG be installed within the LVCPM/SLATE system and integrated with the aircraft to more closely replicate real-world performance at altitude.
3. Perform additional radar model refinement work to continue to add fidelity to the performance of the radar model. Live radar and simulation model differences are exposed when constructive threats are displayed together with live aircraft radar tracks in the cockpit. The effort would require inputs from aircrew subject matter experts and validation of the updates. Testing and validation of the radar model updates should be conducted with a more capable radar bench system, such as that which may be available at OEM. The intent of this research would be to add additional fidelity and realistic performance degradation to the radar model.
4. Perform similar fidelity testing during test fights to address previously reported limitations with the static FITTL lab test environment system. Test flights would improve data comparing live sensors to simulation models and provide the look-down view needed to improve evaluation of TPOD air-to-ground testing trials.

ACKNOWLEDGEMENTS

The authors of this paper wish to thank the Air Force Research Laboratory Warfighter Interactions and Readiness division for providing the resources to conduct this study.

REFERENCES

Call, L. & Lechner, R., (2018). *Advancing the State-of-the-Art in Airborne LVC Training.* Interservice/Industry Training, Simulation, and Education Conference proceedings, Paper 18249.

Kirwan, B., Evans, A., Donohoe, L., Kilner, A., Lamoureux, Atkinson, T., & MacKendrick, H. (1997). *Human Factors in the ATM System Design Life Cycle.* FAA/Eurocontrol ATM R&D Seminar, Paris, France. Internet source.

Lechner, R. & Huether, C., (2008). *Integrated Live Virtual Constructive Technologies Applied to Tactical Aviation Training*. Interservice/Industry Training, Simulation, and Education Conference proceedings, Paper 8147.

Lechner, R. & Wokurka, J., (2010). *Advancements of Integrated LVC Applied to Tactical Aviation Aircraft Training.* Interservice/Industry Training, Simulation, and Education Conference proceedings, Paper 10152.

Lechner, R. & Schwering, J. (2012), *Implementing Integrated LVC for Naval Aviation Training.* Interservice/Industry Training, Simulation, and Education Conference proceedings, Paper 12073.