Creating Effective LVC Training with Augmented Reality

Nathan Jones Koren Odermann	Peter Squire	Adrienne Read Natalie Steinhauser
MARCORSYCOM PM TRASYS	Office of Navy Research	NAWCTSD
Orlando, FL	Arlington, VA	Orlando, FL
Nathan.Jones1@usmc.mil Koren.Odermann@usmc.mil	Peter.Squire@navy.mil	Adrienne.Read@navy.mil Natalie.Steinhauser@navy.mil

ABSTRACT

While live training is often the preferred method for Marines, Force-On-Force (FoF) training exercises lack the visual cues necessary to effectively train Call for Fire (CFF), Close Air Support (CAS), and other engagements. Imagination has been the method of training and feedback on live ranges for decades when weapons and platforms were not available or limited by costs. Augmented Reality (AR) training systems now offer the opportunity to provide realistic visuals of virtual and/or constructive entities and engagements on the live range.

To ensure that AR technology can be utilized to support FoF training, an assessment of the Augmented Immersive Team Trainer (AITT), an AR training system, was conducted to determine how well AITT supports specified training objectives. The AITT system was developed by Office of Naval Research (ONR) and transitioned to Program Manager Training Systems (PM TRASYS). The program office technical assessment team utilized a task and attribute based approach to assess the simulator and simulation on both the activities an individual is required to do in the performance of a specific job (i.e., tasks) and the fidelity the device is required to provide to support that performance (i.e., attributes).

This paper describes the AITT technology and the assessment conducted to support training objectives. In addition, the paper discusses potential for AR technology to enhance the live component in Live-Virtual-Constructive (LVC) environments. This paper provides discussion points based on lessons learned and required development for AR systems being developed to enable an effective LVC training solution.

ABOUT THE AUTHORS

Nathan Jones is the Manpower, Personnel, and Training Technical Support Lead at MCSC PM TRASYS. Mr. Jones has over 15 years of experience in human performance, human-systems integration, assessments, and acquisition program support. He is responsible for overseeing PM TRASYS' training front end analyses, Verification, Validation & Accreditations (VV&As), training effectiveness evaluations (TEEs), and training domain expertise for PM TRASYS.

Koren Odermann is the Lead Program Analyst for Team Lead Range Training System at MCSC, PMTRASYS. Koren Odermann has 10 years in Acquisition Program Management. Koren Odermann is the project manager for the transitioned project Augmented Immersive Team Trainer (AITT), and the development of the new research and development project Mobile Fire Support Trainer (MFST) which leverages the technology from AITT into a new producible training system.

Peter Squire is the Program Officer for Human Performance, Training, and Education (HPT&E), at ONR Code 30. As the HPT&E Program Officer, he manages a broad portfolio of research projects, basic research through advanced technology development efforts spanning a variety of research domains including psychology, physiology, education, and computer science.

Adrienne Read is a Research Psychologist at NAWCTSD. Mrs. Read conducts analyses on naval training and instructional systems to determine effectiveness of design and content delivery. During the analysis of this paper, Mrs. Read was contractor support to PM TRASYS as an employee of TAPE, LLC. Mrs. Read has extensive academic training in psychology and program evaluation and professional experience in curriculum development and instruction.

Natalie Steinhauser is a Senior Research Psychologist at the Naval Air Warfare Center Training Systems Division. She has 11 years of experience with the Navy in experimental design, training effectiveness evaluations, intelligent tutoring systems & adaptive training, usability assessment, and decision-making research. Ms. Steinhauser holds a M.S. in Modeling and Simulation from the University of Central Florida, and a B.S. in Human Factors/Applied Psychology from Embry-Riddle Aeronautical University.

Creating Effective LVC Training with Augmented Reality

Nathan Jones
Koren Odermann
MARCORSYCOM PM TRASYS
Orlando, FL
Nathan.Jones1@usmc.mil
Koren.Odermann@usmc.mil

Peter Squire
Office of Navy Research
Arlington, VA
Peter.Squire@navy.mil

Adrienne Read
Natalie Steinhauser
NAWCTSD
Orlando, FL
Adrienne.Read@navy.mil
Natalie.Steinhauser@navy.mil

GROUND SIMULATION CAPABILITY GAP

One of the biggest challenges of employing effective Live, Virtual, Constructive (LVC) training is providing the live participants with realistic visuals of and interactions with virtual and constructive entities. Ground units particularly suffer from this shortcoming in LVC training exercises. For over half a decade, there has been a strong call from leadership within the infantry community to transform the way ground combat elements train. In order for our ground units to receive the same level of training achievable by other communities via their simulators and simulations, the simulators and simulations for ground units need to reside where ground units train—real-world terrain. There are still many challenges to overcome to fully support and complement live ground training with simulation; however, Augmented Reality (AR) is one possible tool to support synthetic blended LVC training with accurate physics and algorithms to address infantry squad training shortfalls.

DEVELOPMENT OF AUGMENTED REALITY TECHNOLOGY

Augmented Reality inserts virtual elements on top of a real-world view. The classic example of augmented reality is the yellow first-down line in a televised football game. The line is not actually on the playing field, but rather overlaid on top of a live scene. AR technology is the enabling capability for the Augmented Immersive Team Trainer (AITT).

AR is part of a larger Reality-Virtuality (RV) continuum (Milgram et. al 1995). On one side is the real environment.. At the opposite end of this continuum is the virtual environment which is a fully synthetic world. A Mixed Reality (MR) environment includes elements from both the real-world and virtual environments and cuts across much of the RV continuum. As you enhance the real environment with computer generated information, you create an Augmented Reality (AR).

The Office of Naval Research (ONR) has been funding AR research for over 20 years. Early research in AR performed by Steve Feiner led to the development of the Mobile AR System (MARS) (Hollerer, 1999). Feiner and Naval Research Lab (NRL) extended this work and developed the Battlefield AR System (BARS™) (Julier et al. 2000a, b; Livingston et al. 2002). The BARS™ prototype went beyond the MARS campus touring systems and examined how multiple mobile AR users could cooperate with each other as well as with those located in a combat operation center.

Another program that furthered AR progress was the result of the work performed in the Future Immersive Training Environment (FITE) Joint Capability Technology Demonstration (JCTD). The FITE JCTD was a two-year program (2009-2011) that sought to advance simulation technologies to support small unit training. There were two spirals associated with the development of the FITE JCTD. Spiral 1 focused on developing and demonstrating individual immersive virtual reality technologies, while Spiral 2 focused on mixed reality and AR technologies (Muller, 2010). The two key challenges and drivers associated with Spiral 2 were precision mobile AR tracking and rendering. Highly accurate and precise tracking algorithms are necessary to ensure that virtual objects do not move, even when the user is moving or looking around. Once virtual objects are accurately placed and stable within a real-environment, it is necessary to make sure objects look as realistic as possible and match user expectations and environmental conditions. These initial technical challenges were developed in Spiral 2, but carried over into ONR's most recent AR program - AITT.

DEVELOPMENT APPROACH

In 2009, the Naval Research Advisory Committee (NRAC) released the report on Immersive Simulation for Marine Corps Small Unit Training. According to that report, virtual simulations have demonstrated effectiveness for training. They state: "It is believed virtual simulation, with proven effectiveness in training Airmen and Sailors, can improve the safety and effectiveness of Marines and Soldiers by aiding them in developing complex and intuitive decision skills under conditions of stress." However, not all training technology was ready. Specifically, it was determined that technologies associated with AR must be further developed to support the scale and footprint required for United States Marine Corps (USMC) small unit training (Naval Research Advisory Committee, 2009).

The following year, the Marine Requirements Oversight Counsel (MROC) approved the Squad Immersive Training Environment (SITE) Initial Capabilities Development Document (ICD). As defined in the ICD, "SITE is envisioned as an integrating construct or 'toolkit' of LVC training capabilities used to significantly improve infantry squad operational readiness and squad leader tactical decision-making skills." The ICD lists several capability gaps that included the capability to provide realistic battlefield effects to set the conditions for maneuver. At that time, there was limited capability to impose realistic effects and reactions to the employment of fires in order to enable the trainee to accurately assess conditions for maneuver in a training environment.

In 2012, the Marine Corps updated their Science and Technology (S&T) Strategic Plan – (Head Quarters Marine Corps, 2012). This plan established priorities and guidance to support the S&T investment necessary for the future force. Within that document, Science and Technology Objectives (STOs) were outlined to describe the overall S&T requirement. Training and Education is one of the functional areas described in the Strategic Plan which included Warrior Simulation and Warrior Decision-Making. The Warrior Simulation STO calls for the development of simulation capabilities that mimic the live environment (e.g., real-time effects, engagement of senses). The Warrior Decision-Making STO describes the need for the accelerated development of cognitive, relational, and perceptual skills for improving decision-making in complex environments (e.g., expertise, adaptability, problem solving). Together, these STOs describe the need to support small units' ability to train as they would fight, and describe capabilities that support the development of critical reasoning and other decision making skills across a range of military operations.

One of the key gaps for Infantry simulation was the ability to see simulated battlefield effects, such as munitions effects, overlaid on the real-world. For example, during Force-On-Force (FOF) exercises, the ability to see where a simulated munition lands provides the necessary visual cues to support adjusting fires. Without these visual cues, Marines are left to their imagination and hand waving, which limits the training capability for United States Marine Corps (USMC). The AITT effort sought to address this training gap and requirement using AR as a proof of concept demonstration.

AUGMENTED IMMERSIVE TEAM TRAINER

ONR began developing the AITT in Fiscal Year 2011 (FY11) to provide a capability for providing realistic battlefield effects to set the conditions for maneuver. There were three primary performers: (1) SRI International was selected to develop the hardware and tracking; (2) Lockheed Martin was selected to develop simulation / rendering and to lead the integration of AITT components; (3) UCF was selected to conduct the usability evaluations and training assessments. Major demonstrations were held each year through FY16; notable demonstrations took place at Marine Corps Base Camp Pendleton and Marine Corps Base Quantico [Range 7, 15, Lejeune Field, and Medal of Honor Golf Course]. Additional demonstrations occurred at Fort Bliss (Network Integration Evaluation 15.1 – Army Funded), Army Expeditionary Warrior Experiment, Bold Quest 14.2. Ultimately, the success of these demonstrations led to a transition to Marine Corps Training and

Figure 1: AITT System

Education Command (TECOM) and Program Manager Training Systems (PM TRASYS) in FY16.

AITT combines the real-world scene with realistic virtual elements (e.g., mortar and artillery weapons effects, fixed and rotary wing aircraft) and targets (e.g., enemy personnel, tanks, or buildings). A Commercial Off-the-Shelf (COTS) video see-through Head-Mounted Display (HMD) allows the user to see the resulting images overlaid on the live terrain. Figure 1 shows the view that the trainee sees through the HMD. In the scene, the trainee sees the real life live range with virtual tanks added into the scene. The virtual tanks are highlighted with yellow arrows. These virtual entities follow the terrain the same as a live vehicle would. As the trainee turns his head to scan the scene, the virtual entities stay in the appropriate place in the HMD view.

Figure 2: Virtual tanks overlaid on live range at Camp Pendleton Wiskey Range viewed via HMD.

The AITT system utilizes a camera to process the terrain and real external environment. The system adds the virtual entities to the scene and presents the combined picture to the trainee. The AITT system enacts effects on the targets, and allows interfacing with other Marine Corps systems such as Instrumented-Tactical Engagement Simulation System (I-TESS) or Deployable Virtual Training Environment (DVTE) Combined Arms Network (CAN).

TECHNOLOGY READINESS ASSESSMENT

When AITT transitioned to TECOM and PM TRASYS in FY16, close to six years had passed since the Marine Requirement Oversight Counsel approved the SITE ICD that outlined the capability gap to provide realistic battlefield effects to set the conditions for maneuver. To ensure the system addressed the capability gap and determine what research and development was necessary, TECOM and PM TRASYS conducted a readiness assessment in terms of ability to support training. If the simulator was found to be lacking the necessary functionality and fidelity in comparison with live training, the benefit of the simulation training would be lost. The primary goal of this evaluation was to determine the AITT's capability to support training requirements on tasks related to its intended use to enable Call For Fire (CFF) in FOF training. Due to large interest and perceived potential to also support Close Air Support (CAS) operations, an assessment to support Joint Terminal Attack Controller (JTAC) training was conducted at the same time. The evaluation scope did not include assessing the level of learning achieved by trainees.

Methodology

The team utilized a systematic process of assessing the capability and attribute fidelity of AITT to support independent tasks derived from Training and Readiness (T&R) Events. T&R Events are training requirements identified in T&R Manuals for each Military Occupational Specialty (MOS). T&R Manuals establish the training standards, regulations, and policies for training each respective MOS. Satisfaction of T&R Events is utilized to measure unit readiness.

The team derived tasks to be evaluated from the following publications and T&R manuals: NAVMC 3500.42B Tactical Air Control Party (TACP) T&R Manual and JTAC Memorandum of Agreement (MOA), NAVMC 3500.7B

Artillery T&R Manual, NAVMC 3500.44B Infantry T&R Manual, and NAVMC 3500.23A Air Naval Gunfire Liaison Company (ANGLICO) T&R Manual. The T&R Manual review identified one hundred and sixteen (116) T&R Events to be evaluated based on perceived ability to support. Ninety four (94) events pertained to the Artillery/Mortar user community and twenty two (22) T&R Events pertained to the JTAC user community.

The evaluation was conducted at Camp Pendleton in March 2016. The assessment team was led by PM TRASYS and TECOM with support from General Dynamics Mission Systems, SRI International, and Lockheed Martin. Ten subject matter experts (SMEs), six Forward Observer (FO) SMEs from the Artillery/Mortar community and four JTAC SMEs, participated in the capability rating portion of the assessment. . JTAC-qualified SMEs evaluated all tasks (CFF and CAS) used for the evaluation. Artillery and Mortar SME participants only evaluated CFF tasks..

The set of training system attributes required to effectively support task performance was established prior to the evaluation and utilized to measure if the fidelity and functionality of AITT supported each task. Seventeen attributes were evaluated for each task:: physical appearance, tactile feel, physical environment, haptics, systems response, sound bearing, environmental noise, verbal communication, audible signals, static visuals, active visuals, resolution, visual depth, light levels, motion awareness, horizontal field-of-view, and vertical field-of-view. Qualitative data was obtained by collecting feedback from the evaluation participants and the system operators during the assessment and through structured interviews during After Action Reviews (AARs) after the evaluation was completed.

Task Analysis

The T&R Events were analyzed at the independent task level for each user community. A task analysis was conducted on the identified 116 T&R Events to derive the independent tasks required to complete all identified T&R Events. Specific task statements were written in a standard format based on the tasks identified by the SMEs. The standardized format consisted of an action verb, an object, and possible qualifiers (where applicable), which ensures the intended action is clearly described, is specific, and can be measured.

An example of the derivation process can be demonstrated using the T&R performance step: “perform battle damage assessment.” The performance step actually represents a large number of potential independent tasks that may need to be performed by a trainee. Therefore, the team identified each one of those options, for example: “conduct BDA for time fuze;” “conduct BDA for quick fuze;” “conduct BDA for delay fuze;” “conduct BDA for 81mm mortar;” “conduct BDA for 120mm mortar;” and so on.

Two hundred, ninety-one (291) independent tasks were derived from the 116 T&R Events. There were two reviews of the task list and reductions in number of these tasks to be evaluated. The first reduction excluded one hundred and twenty (120) tasks due to redundancies, scope of the assessment, or updates to T&R Manuals. The second reduction was applied due to known limitations of the current AITT. Limitations included specific ordinance not yet supported, specific target types/equipment not yet supported, specific weather effects not yet supported, or unsupported night operations. Seventy eight (78) tasks were excluded due to system limitations. After exclusions, only ninety-three (93) tasks remained for evaluation.

Scoring for T&R Events

The evaluation of AITT’s ability to support a T&R Event was done through a three-step process.

1.) Rate the criticality of each attribute for enabling each task on a 1 to 5 scale. A rating of one means the attribute is irrelevant and contributes nothing. A rating of five means the attribute is absolutely critical and the task cannot be executed without this attribute. This rating was conducted prior to the capability assessment of the AITT system. The SMEs rated and reached concurrence on how critical the presence of each of these attributes was to complete the particular tasks being analyzed using the rating scale. For example, when considering the task of “operating target location equipment,” the SMEs determined the Appearance (physical properties) and the Tactile Feel (touch sensation) of the simulated equipment required high fidelity and were critical attributes (both rated at a 4). However, for this same task, the Auditory attributes (Environmental & Battle Sound, Sound Bearing, Audible Systems Signals, and Verbal Communication) were rated as Irrelevant because these attributes contribute nothing to executing the task of operating target location equipment.

2.) SMEs performed tasks with AITT. The evaluation team utilized written scripts for each vignette to minimize any potential bias on outcomes. SMEs rated the capability of each attribute to support each task on a 1 to 5 scale. A rating of one means the system is completely incapable of providing the specified attribute and cannot support task execution. A rating of five means the system is fully capable of providing the specified attribute with no departure from realism.

3.) The collected ratings of the independent tasks required to complete each T&R event determine an overall Code Training Support (CTS) score for each T&R Event. The CTS score indicates how well the system supports the T&R Event. CTS is a weighted average of the attribute criticality and capability ratings for all tasks required for the T&R Event. Each task is given a score based on the criticality and capability ratings; these scores are used to calculate CTS scores. For example, if a T&R Event has 3 required tasks, the weighted average of the attribute scores for all 3 tasks is calculated as a single score. Higher weight goes to attributes with higher criticality. If an attribute has a high capability, but a low criticality, the weighted average will be lower. Although consensus in ratings was achieved for the criticality ratings, consensus was not sought on capability ratings. Each SME performed the tasks independently and provided their own ratings. The CTS score is then used to determine if AITT supports the task based on the score definitions. A Level 1 rating indicates no training capability. Level 5 indicates T&R Event can be thoroughly and accurately trained.

RESULTS

There were 96 tasks executed that were derived from 116 T&R Events. Overall, the participants were impressed with the AITT capabilities. The SMEs stated they would implement the AITT as a tool for basic FO training as soon as possible for static location and lane training, target recognition and identification, determining locations, map correlation exercises, selecting appropriate munitions for different target types, and rehearsing scripts for CFF and CAS. With additional capabilities, the SMEs saw potential uses for learning the effects of munitions that are not seen in live training due to safety issues or cost; as well as integrating with other systems to allow multiple groups to engage in the same simulated exercise spread out across the live terrain. The JTAC SMEs identified that current CFF/CAS trainers, such as the Supporting Arms Virtual Trainer (SAVT), are limited to 260 degree viewing, but the AITT allows for 360 degree field of view.

Qualitative Analysis

Several different results were obtained from the focus group discussions. First, is the HMD image quality. The contrast of light and dark colored images portrayed through the HMD caused loss of resolution, making the projection of the live environment appear blurry. This contrast was dependent on the time of day and the position of the sun (whether or not direct sunlight was present). When it was late in the afternoon or if there was cloud cover, light levels through the AITT had better contrast and were more realistic compared to light levels when viewing with the naked eye. However, when there was direct sunlight (early in the day and at high noon) and when there was no cloud cover, the contrast caused lighting and resolution deficiencies. The ability of HMDs to work under bright condition is a known challenge for mobile outdoor AR.

One of the key enablers to the technology is correlation of the live terrain to the AITT's hosted terrain map. During the assessment, limitations were encountered with the ability for the AITT to maintain accurate terrain correlation when the trainee moved around from a fixed location (team did not measure and determine effective distance trainee could move). This is a major limitation of the S&T development that prevents FOF movement through the range (restricts the trainee to known locations). The AITT system also demonstrated limitations with rendering objects on the live terrain. At times, entities would appear to be floating; the floating was an issue with rendering details. The same rendering details created issues with lack of accurate contrast with the environment, making virtual/constructive entities stand out noticeably from their surroundings depending on the background.

The AITT system limits occlusion to objects that are provided in the AITT terrain model. For example, vegetation is not included in the model. As such, the system does not know when a virtual entity is placed behind a tree or building, and therefore, it is not hidden. Users can manually insert objects (e.g. tree, buildings, etc.) but the process is time consuming. Occlusion issues arose when a simulated object in the distance (such as an armored vehicle), was not appropriately occluded by a closer object in the live environment (such as a tree).

Lastly, there is a potential workload issue for the instructor. The instructor has to input commands based on what the trainee does, as well as role play other positions and control how the entities react. Instructor responsibilities currently include targeting input, scenario building, airborne entity control, rate of fire, and Fire Support Coordination Measures (FSCM).

T&R Events Supported

Improvements in system capabilities for the AITT to enable T&R Events were determined by the identification of the system attributes that failed to support task performance combined with a qualitative assessment with each of the SME participants during data collection and during AARs. A CTS score below Level 3 indicated T&R Events that were not currently supported by AITT. The AITT system supported training on 55 (47%) of the evaluated 116 T&R Events with a CTS Level 5 or Level 4, 25 (22%) received a CTS Level 3, and 35 events (30%) received a CTS score below Level 3. A sampling of T&R Events with the highest scores (Levels 4 & 5) include: Conduct an immediate suppression mission; Conduct an illumination mission; Integrate SEAD during the execution of CAS missions; Conduct a quick smoke mission; Conduct a danger close fire mission; Adjust fire with the laser range finder; Conduct a Fire for Effect (FFE) mission; Conduct two fire missions simultaneously; Locate a target; Conduct a destruction mission; Adjust Final Protective Fires (FPF); Conduct a naval surface fire support (NSFS) mission; Conduct Fire Support Team (FiST) operations; Determine target location via coupled GPS/LRF system; Provide situation update to CAS aircraft; and a few more.

Perhaps just as important are the 35 T&R Events with scores below Level 3, which indicate the T&R Events that AITT does not currently have the capability to support during training. A sampling of these include: Adjust fire with the laser range finder; Conduct a coordinated illumination mission; Operate organic digital aided CAS/Fires systems; Control an assault support platform into a marked LZ; Conduct simultaneous missions with NSFS; Conduct an adjust fire mission; Execute Type 1, Type 2, or Type 3 terminal attack control procedures; Conduct a danger close fire mission with NSFS; Conduct a mortar precision registration; Conduct an immediate smoke mission; Conduct a naval gunfire (NGF) coordinated illumination mission; Adjust final protective fires; Conduct a coordinated illumination mission; Conduct a Suppression of Enemy Air Defense (SEAD) fire mission; Conduct an immediate smoke mission; Conduct a precision registration, quick and time; and a few more.

The underlying issues across T&R Events that need to be addressed include: inaccurate munitions, limited battle damage and environmental effects, difficulty acquiring aircraft in flight, contrast of entities with surroundings, correlating live and synthetic terrain (elevation and topography) with visual system (land marking), sound inadequate and non-directional, inability to traverse terrain, targeting capabilities (grid, polar, shift from known point), target placement accuracy, lack of control over airborne entities, and lack of control over firing rate (one round at a time).

Lesson Learned: User Populations

Material development capability prioritization should be dictated by the training needs of user communities. The intended primary training audience and their specific task performance needs should be addressed first, which requires recognizing there is a unique difference in user community needs. After the primary audience needs are satisfied, developers should not assume the developed system will be sufficient for other training audiences. Each MOS with a significant vested interest in the training system should be evaluated as independent audiences. A user group of multiple MOSs should not be treated as one homogeneous group until test measures statistically support it. In this case, there was significant enough difference to treat the future development of CFF and CAS capabilities as different development requirements.

The AITT S&T project was developed specifically for the Mortar/Artillery FOs. A more in-depth analysis of the results revealed some significant differences in ratings between the user community groups indicating AITT S&T system cannot be utilized for JTAC without more development to support their specific needs. The results influenced the development decision for the next phase of research and development tasks to support JTAC training for the new Mobile Fire Support Trainer (MFST).

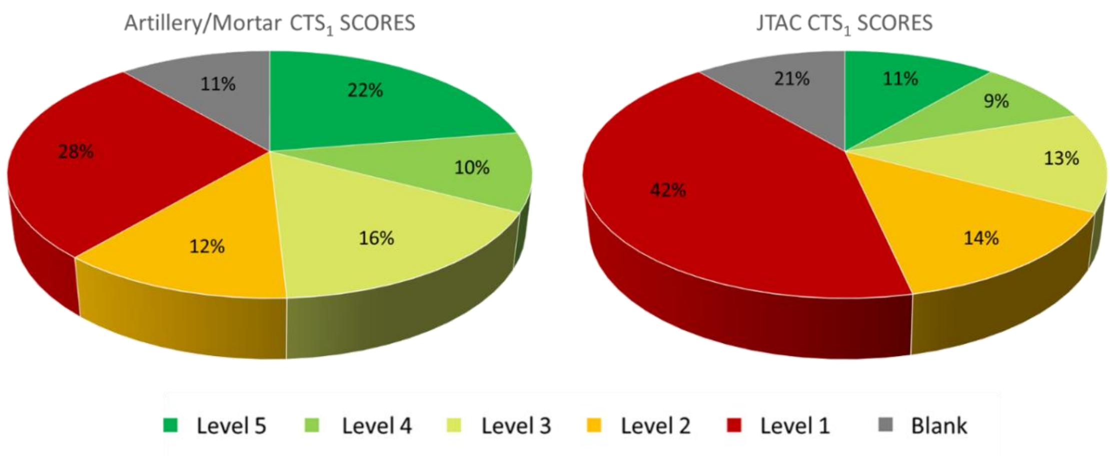

Figure 3: CTS Score Comparison between Artillery/Mortar and JTAC

The T&R Events with the greatest differences in CTS scores include (but are not limited to): Locate a target using digital systems; Locate a target; Conduct an illumination mission; Employ supporting arms; Conduct an adjust fire mission; Act as an observer for Close Air Support (CAS); Conduct a naval surface fire support (NSFS) mission; Conduct an illumination mission with naval surface fire support (NSFS); Call for indirect fire using the grid method; Conduct a Coordinate Illumination Mission; Call for indirect fire using the shift from a known point method.

The differences in the ratings of the two communities should not be overlooked during system design and upgrades. The difference in CTS scores depended on the raters' competency groups (Artillery/Mortar versus JTAC); therefore, a Mann-Whitney U Test was performed to determine if the difference was statistically significant differences between groups. The Mann-Whitney U was based on the scores for tasks executed by both communities deemed critical to the execution of each T&R Event. The results for the Mann-Whitney U Test are depicted in Table 1. The results indicate there is a statistically significant difference in CTS scores between the Artillery/Mortar and JTAC competency groups based on a significance level of alpha at 0.05 ($p = 0.0203$).

Table 1: CTS Mann-Whitney U Test Data Output

Mann-Whitney U	Standard Error	Standardized Test Statistic	*df*	*p*-Value	95% Confidence Interval	
					Lower	Upper
4311.5	42.1515	2.3201	102	0.0203	0.4	4.3

The significant differences between user communities is an important finding given the tasks under assessment originate from different T&R Manuals. When the tasks rated are derived from T&R Events from the Artillery/Mortar communities (i.e., Artillery, Infantry, and ANGLICO T&R Manuals), that user community's ratings are more valid than the JTAC community's ratings for those specific tasks. While both communities may perform tasks perceived to be the same in execution of their T&R Events, there is a measured difference in how they execute the task and their expectations of how the simulator and simulation should support the task with necessary stimuli.

Kappa was calculated for each group's lowest scoring attributes to determine if the SMEs were in agreement. Kappa goes from zero (no agreement) to one (perfect agreement). Table 2 includes the kappa statistics for the attributes rated three or below by each SMEs group across all tasks. Artillery/Mortar SMEs had strong agreement that their worst scoring attributes were deficient. Therefore, confidence was obtained by the assessment team that these attributes are in fact deficient for supporting Artillery/Mortar tasks. Whereas, the JTAC SMEs lacked agreement for determining attributes were deficient. Therefore, in order to address system improvement needs for JTAC, additional analysis and retesting will be necessary.

Table 2: SME Agreement for Attributes Rated Three or Below

	Attribute	Kappa Statistic	Level of Agreement
Artillery/Mortar	Static Visuals	0.92	Almost Perfect Agreement
	Resolution	0.92	Almost Perfect Agreement
	System Response & Interaction	0.79	Substantial Agreement
	Active Visuals	0.62	Substantial Agreement
JTAC	Resolution	0.10	Slight Agreement
	Active Visuals	0.19	Slight Agreement
	Depth Perception	0.21	Fair Agreement
	Light Levels & Celestials	0.21	Fair Agreement
	Sound Bearing	0.24	Fair Agreement

*Kappa scores presented are only for the lowest attribute. It does not address agreement for higher scoring attributes.

Limitations

The results presented are based solely on the 93 tasks evaluated. Future evaluations will need to include the 78 tasks that were not tested due to current system limitations, which will affect 67 (58%) of the T&R Events. Additionally, data collection was limited to the number of available SMEs and was bound by a one-week timeframe for data collection.

TECHNOLOGY IMPROVEMENT DISCUSSION

In order for the AITT system, or any AR system, to support CFF or CAS training, the following engineering issues need to be considered. This evaluation was supportive for the systems requirements review process for the development of the MFST, the follow on project to AITT. Additionally, a new ONR effort, Warfighter Augmented Reality (WAR), has already started to address several of the challenges.

Maneuvering in Live Environment

In its current tested stage, AITT is limited in its ability to support LVC training on the live range. The AITT demonstrated successfully that it can initially be utilized for stationary training. The technology limits the LVC capability to fixed known locations. FOF is not yet fully achievable as free-play or traversing on open range (without pre-established observation points that can be loaded into the system prior to the start of the FOF exercise). Continued development by the Government and industry partners is needed to effectively enable LVC supported CFF and CAS on the live range. The WAR effort is seeking to increase the freedom of movement. There is a tradeoff between the accuracy of the system, and how much gear is required to wear. In addition, there are newer techniques that are being investigated (e.g. Dynamic Landmark matching) to address the accuracy and maneuverability for the AR system. The new development of MFST will expand the ability of the user to maneuver within the terrain.

Light Levels

Visual light levels will need to be controllable for the live environment and for the virtual/constructive entities, and the responsivity to light requires adjustment to limit the negative impact to projected images. Likewise, the interaction of light with the virtual/constructive entities requires interactive details. The lack of realistic interaction made it difficult to acquire aircraft in flight. The simulated aircraft were missing cues that would exist in real life, such as the sun reflecting off the aircraft. While there are software solutions to address the light levels, many of the challenges are also hardware based. Specifically, the ability to go towards optical-see-through requires much more work to solve the associated technological challenges. This is an opportunity, specifically for industry and academia, to provide solutions.

Terrain Mapping

The terrain map hosted by the AITT does not include a level of detail down to individual objects to correct occlusion issues. While the system was able to obscure virtual objects based on topographical terrain, it could not differentiate some physical objects on the terrain. Virtual objects appear to be placed on top of other live objects (trees, vehicles, buildings) vice being appropriately integrated into the scenery. Recent work has addressed this issue by allowing Marines to collect and process their own terrain models (McAlinden, et al., 2015). In contrast, to terrain models that only provide elevation data, new techniques support terrain models that include objects. This works well for objects that will not move (e.g. trees, buildings). However, this does not support dynamic occlusion, when people or vehicles are moving in the AR view. More work is necessary to be able to support that capability.

User Interfaces

The issues associated with the I/O primarily involve the system's user interface. Streamlining the process for data input, targeting, and scenario control will decrease the workload required to operate the AITT, allowing the I/O to focus more on the trainee and the training exercise. Beyond all the duties for the I/O, the I/O is limited in ability to see what the trainee sees. Incorporating live feeds and AAR capabilities such as video recording or playback would provide the I/O the trainee view information for more effective training.

Man-worn Fit

The bulkiness, size, and weight on the head for HMDs, batteries, and cables is an issue for how long trainees can wear the equipment during training scenario as well as limiting maneuvering on a range. Continued reduction in size of cameras, batteries, and man-worn computers is vital to the practicality of utilizing AR on the range.

CONCLUSION

The AITT demonstrated a capability to support T&R Events for CFF in FOF training. It is currently able to support lane training and schoolhouse curriculum in pre-determined, static locations or FOF exercises with stationary FOs. With additional research and development, it should be able to support FOF free play exercises as well as JTAC training, which was beyond the original requirement. The attributes rated as most deficient between user communities included many of the same attribute categories, but they held different priorities within each group. This has been taken into consideration in development so that priority for engineering efforts is assigned according to priority users. The program management office will continue to work with industry partners to remedy the known limitations to enable FOF training as well as focus on system hardening, form and fit, and software improvements to expand entity modeling and system functionality.

The implementation of the recommended upgrades will improve MFST's and WAR's ability to support execution of T&R Events and enable live participants in an LVC training exercise to directly and effectively interface with virtual and/or constructive entities. AR technology development should be focused on the issues concerning scene rendering details, terrain correlation, and ability to use anywhere on a live range.

Overall, the assessment demonstrated the capability for AR technology to become the interface between the live and virtual/constructive worlds in live training exercises for CFF, CAS, and FOF applications. Their primary benefit is presenting live, virtual, and constructive entities in the same picture to a trainee. The manipulation of that environment may be controlled by the AR system's computer or be provided by other federated or interoperable sources in an LVC environment. No other technology offers the ability to truly see virtual and constructive entities in the live environment, thereby creating a realistic LVC training environment.

ACKNOWLEDGEMENTS

The authors want to acknowledge the efforts provided by the entire assessment team: Capt Michael Eady and Dave Dunfee from TECOM; Quy Nguyen and Chuck Peabody from PM TRASYS; Rob Leavitt and Kevin Mullally from General Dynamics; and Sean Cullen and Mikhail Sizintsev from the Lockheed Martin/SRI International AITT development team. As well, we thank Luis Velazquez and Aerial Kreiner of the IITSEC subcommittee for guidance in developing this paper.

REFERENCES

Baillot, Y. (2006). System and method to display maintenance and operational instructions of an apparatus using augmented reality. U.S. Patent Application 11/441,241.

Defense Acquisition Guidebook. (2015). https://dag.dau.mil

Director, Joint Staff J8, *Joint Fire Support Executive Steering Committee (JFS ESC) Action Plan Memorandum of Agreement (AP MOA) 2004-01, Joint Terminal Attack Controller (JTAC) (Ground)*, May 2015

Goldiez, B., Livingston, M. A., Dawson, J., Brown, D., Hancok, P., Baillot, Y., & Julier, S. J. (2004). Advancing human centered augmented reality research. University of Central Florida, Orlando, FL.

Head Quarters Marine Corps. (2012). Marine Corps S&T Strategic Plan: Leading edge technology for the Marines of tomorrow.

Höllerer, T., Feiner, S., Terauchi, T., Rashid, G., & Hallaway, D. (1999). Exploring MARS: developing indoor and outdoor user interfaces to a mobile augmented reality system. Computers & Graphics, 23(6), 779-785. doi:10.1016/S0097-8493(99)00103-X.

McAlinden, M., Suma, E., Grechkin, T., Enloe, M., (2015) Procedural Reconstruction of Simulation Terrain Using Drones; I/ITSEC

McHugh, M. (2012). Interrater reliability: the kappa statistic. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995, December). Augmented reality: A class of displays on the reality-virtuality continuum. In Photonics for industrial applications (pp. 282-292). International Society for Optics and Photonics.

Muller, P (2010). The Future Immersive Training Environment (FITE) JTCD: Improving Readiness Through Innovation; I/ITSEC

Naval Air Warfare Center Training Systems Division (NAWCTSD), *SAVT Systematic Team Assessment of Readiness Training (START) Evaluation Report*, July 2013

Naval Research Advisory Committee (2009). Immersive Simulation for Marine Corps Small Unit Training.

Navy Marine Corps Instruction (NAVMC) 3500.7A, *Artillery Training and Readiness Manual (Artillery T&R Manual)*, November 2012

Navy Marine Corps Instruction (NAVMC) 3500.23A, *Air Naval Gunfire Liaison Company Training and Readiness Manual (ANGLICO T&R Manual)*, October 2012

Navy Marine Corps Instruction (NAVMC) 3500.42B, *Tactical Air Control Party Training and Readiness Manual (TACP T&R Manual)*, May 2014

Navy Marine Corps Instruction (NAVMC) 3500.44B, *Infantry Training and Readiness Manual (Infantry T&R Manual)*, August 2013

Rosenblum L, Feiner S, Julier S, Swan JE, II, Livingston M (2012). The development of mobile augmented reality. In: Dill J, Earnshaw R, Kasik D, Vince J, Wong PC (eds) Expanding the Frontiers of Visual Analytics and Visualization. Springer London (pp 431-448). doi:10.1007/978-1-4471-2804-5_24.