Optimum Dismounted Soldier Training Experience: Live or Virtual?

Emilie A. Reitz
General Dynamics – Information Technology
Suffolk, VA
emilie.a.reitz.ctr@mail.mil

Robert Richards
Alion Science and Technology
Suffolk, VA
rrichards@alionscience.com

ABSTRACT

A fire team mounts the back stairs of a two story duplex and is quickly picked off by insurgents firing from inside the top floor. With four men down, the squad leader shakes his head and takes off his helmet. Thankfully, this occurred in an immersive environment at Fort Benning, Ga., during a capability demonstration and assessment hosted by the U.S. Army Maneuver Center of Excellence, Maneuver Battle Lab and Joint Staff J6 as part of Army Expeditionary Warrior Experiment Bold Quest 2012-2 (AEWE-BQ12-2) (Reitz & Seavey, 2012). The squad leader asks the exercise controller to stop and send it to the After Action Review screen. His unit regroups around him, and they talk through what just occurred.

AEWE-BQ12-2 allowed for an assessment of the quality of training transfer from virtual training capabilities to live mission execution. One area assessed was the impact of training capabilities on situational awareness, understanding and small unit readiness. Over three weeks, four squads performed area reconnaissance, cordon and search, and attack missions. Three squads trained in a digitized McKenna Military Operations in Urban Terrain (MOUT) virtual environment for three days before performing each live mission. The control squad utilized the live MOUT environment for both training and execution of their missions. The three experimental squads (N=34) performed at the same level as the control squad in situational awareness as measured by Situational Awareness Rating Tool (SART) (Taylor, 1990). Inter- and intra-squad communications increased significantly in the experimental squads, compared to a control.

Future austere budgetary environments lead military leaders to look to virtual training environments to hone combat proficiency within the force. This paper details results of the experiment, and outlines implications of those results on balancing live and virtual training capabilities to enhance situational understanding and small unit readiness.

ABOUT THE AUTHORS

Emilie A. Reitz is a Research Analyst at General Dynamics Information Technology. She is currently supporting the Joint Fires Division of Joint Staff J6, Deputy Director for Command and Control Integration (C2I). Her research focuses on integrating joint capabilities into modeling, simulation, and training. She was the research lead for the JCW immersive technology initiative during BQ11; Joint Staff lead Analyst during AEWE-BQ12-2 and acted as a research assistant for the NTSA award-winning Border Hunter initiative. Emilie holds a Master's Degree in international studies.

Robert J. Richards is a Senior Systems Engineer at Alion Science and Technology. He is currently supporting the Joint Fires Division of Joint Staff J6, Deputy Director for Command and Control Integration (C2I). In this capacity he provides support for the integration of Joint / Coalition Live and Virtual tactical training capabilities, to include application of human factors assessment of Joint training and advanced simulation capabilities to improve human decision making and effectiveness. He was the lead US / Allied integrator for Bold Quest during AEWE-BQ12-2. Bob has a Master's Degree in Aeronautical Science.

Optimum Dismounted Soldier Training Experience: Live or Virtual?

Emilie A. Reitz
General Dynamics – Information Technology
Suffolk, VA
emilie.a.reitz.ctr@mail.mil

Robert Richards
Alion Science and Technology
Suffolk, VA
rrichards@alionscience.com

INTRODUCTION

A fire team mounts the back stairs of a two story duplex and is quickly picked off by insurgents firing from inside the top floor. The building was a key objective for their mission, to provide cover for another maneuvering unit, but the squad has now completely lost combat effectiveness; the squad leader's focus has shifted to maintaining cover, recovering his downed men, retreating from the building, and medivac. He stands up from his crouching position, and signals to the trainer to end the scenario. Thankfully, this occurred in an immersive environment at Fort Benning, Ga., during a capability demonstration and assessment hosted by the U.S. Army Maneuver Center of Excellence, Maneuver Battle Lab and Joint Staff J6 as part of Army Expeditionary Warrior Experiment Bold Quest 2012-2 (AEWE-BQ12-2) (Reitz & Seavey, 2012). The squad leader asks the exercise controller to stop and send the ground truth capture of the scenario to the After Action Review screen. His unit regroups around him. They talk through what just occurred, and discuss ways to avoid a similar outcome in the future.

AEWE-BQ12-2 allowed for an assessment of the quality of training transfer from virtual training capabilities to live mission execution. One area assessed was the impact of training capabilities on situational awareness (SA), situational understanding (SU) and small unit readiness. Over a three week training period, four dismounted infantry squads (two U.S. Army, one U.S. Marine Corps and one Canadian Army) performed area reconnaissance, cordon and search, and attack missions. Three squads trained in a digitized McKenna Military Operations in Urban Terrain (MOUT) virtual environment for three days on each mission before performing it live at McKenna MOUT. This paper details results of the experiment, and outlines implications of those results on balancing live and virtual training capabilities to enhance situational understanding and small unit readiness.

IMPROVING SMALL UNIT PERFORMANCE

In order to build on and sustain operational and training superiority, the Department of Defense (DOD) must be able to effectively and efficiently prepare future warfighters with increasingly limited fiscal, time, materiel, and personnel resources. Multiple DOD documents provide overlapping, mutually supportive strategies to reach an end state of innovative, adaptive leaders and small units who are able to perform at an enhanced level, make decisions, and accurately read and respond appropriately to threats in their environment (AUSA Torchbearer, 2011; DOD 2010; Department of the Army, 2011; TRADOC, 2010; TRADOC, 2011). These guiding documents point to the need for training solutions that stimulate cognition, intuition, innovation and adaptive thinking, as well as promote complex decision-making skills at all levels.

The technological capacity and andragogical sophistication have continued to develop, as have the understanding of the ‘human factors’ which underpin our training gaps (Wilson, et al, 2007). There is a growing body of research on using virtual environments as preparation to maximize the effective use of real-world training environments (Roman & Brown, 2009; Dorsey, Russell, and White, 2009; Naval Research Advisory Committee, 2009;), to transfer basic knowledge in an efficient and asynchronous manner. Because of the increasing understanding that virtual training can play a key role in force capability development, there remains a need to better understand the interaction between military skills training, virtual exercises to reinforce those skills, and live field training, as well as the potential operational impacts of that training.

In a teaming effort between the Joint Staff (JS) and US Army Training and Doctrine Command (TRADOC)’s Maneuver Center of Excellence (MCoE), AEWE-BQ12-2 provided a live and virtual experimental environment to inform Joint and Coalition capability development and acquisition relative to combat preparedness and identification for dismounted Warfighters.

Historically, AEWE has served to enhance the capabilities and effectiveness of the current force, simultaneously contributing to future force interoperability and effectiveness, and informing Brigade Combat Team modernization efforts. It serves as a venue to provide capability developers, the science and technology community, and industry a repeatable, credible, rigorous and validated operational experiment venue to support Doctrine, Organization, Training, and Leadership concepts and materiel development efforts.

Bold Quest (BQ) is the JS-led Coalition Capability Demonstration and Assessment conceived in 2001. It provides a repeatable mechanism for multi-national, multi-initiative capability development and testing in a Coalition operational context. Focused on both near- and long-term outcomes, BQ investigates materiel and non-materiel developments. BQ11, led by JS Command, Control, Communications and Computers Joint Fires Division, attempted to address weaknesses in human factors associated with combat identification – communication, cognitive biases, stress and more – by assessing a sequence of mutually supportive training initiatives to create enhanced SA, and provide a basis for improved decision-making and target identification/engagement skills. This progressed naturally into the AEWE-BQ12-2 partnership. BQ11 leveraged artifacts from the US Joint Forces Command (USJFCOM) sponsored Future Immersive Training Environment Joint Capability Technology Demonstration Spiral I, which focused on innovative technical and human effectiveness capabilities associated with training ground forces in interoperability, shared decision-making, synchronized SA and coping with current demands on ground forces operating in stressful, often chaotic, environments (USJFCOM, 2010).

Situational Awareness

Improving dismounted soldier situational awareness is key for all force developers. Within the world of combat identification and target engagement, situational awareness is a lynchpin in a soldier’s ability to be effective on the battlefield. Technologies available for use in combat identification and target engagement can be highly accurate, yet human operators must employ the system and then correctly respond to signals received. Both represent activities subject to error, sometimes with grave consequences. Situational awareness is necessary during the response phase, as the human is the one who acts on the system-provided information with a weapon or a call for fires (Greitzer & Andrews, 2010; Shrader, 1982).

FM 3-0, Operations states that SU occurs at the commander (CDR) level and defines it as “The product of applying analysis and judgment to relevant information to determine the relationships among the mission variables to facilitate decision-making” (Department of the Army, 2001, P 7-11). FM 3-0 further defines SA as "Immediate knowledge of the conditions of the operation, constrained geographically and in time" (Department of the Army, 2001, p C-1).

These definitions are too abstract to be directly applied to squad-level operations. Other definitions of SA and SU break the concept into activities that are the building blocks of the more formal FM 3-0 concept. Endsley (1995a, 1995b) proposed three interleaved levels for SA: Level 1: Perception of Elements in Current Situation; Level 2: Comprehension of Current Situation; and Level 3: Projection of Future Status. The process of perceiving the environment, comprehending the activities occurring in the environment, and then mentally modeling that information to inform the next action are activities regularly performed by individuals and small units during field operations. SA then supplements small unit readiness, which was defined for the purpose of AEWE BQ12-2 as the performance level of the squad achieving the desired end state and satisfying the key tasks established a the higher headquarters five-paragraph tactical order, demonstrated through a field assessment.

METHODOLOGY

AEWE-BQ12-2 was a live, prototypical, force-on-force experiment with a primary focus on the soldier and small unit, examining concepts and capabilities for the current and future force across all warfighting functions. This paper focuses on the elements of the AEWE-BQ12-2 experiment illustrated in Figure 1. Those elements began with five days of Advanced SA Training (ASAT); and three days of baseline field assessment. This baseline study provided solid experimental basis for assessment of squad performance changes across three 96-hour use case scenarios, each consisting of a three-day training period followed by a force-on-force field assessment event. The three case scenarios performed were Area Reconnaissance (AR), Cordon and Search (C&S), and Attack (AT).

A two-group design was utilized with one control and one experimental cohort. The control group only received exposure to two non-confounding training capabilities which were considered part of the assumed assessment baseline: ASAT and the Mission Rehearsal Planning System. During the rest of the exercise, they utilized standard Tactics, Techniques and Procedures (TTP) in their training for the mission assigned for that week. A battery of tests and survey apparatus were administered at key points – before the onset of the exercise, after completion of each training period, and after each field assessment. Each week, one squad was removed from the direct force on force assessment due to conflicting assessment requirements.

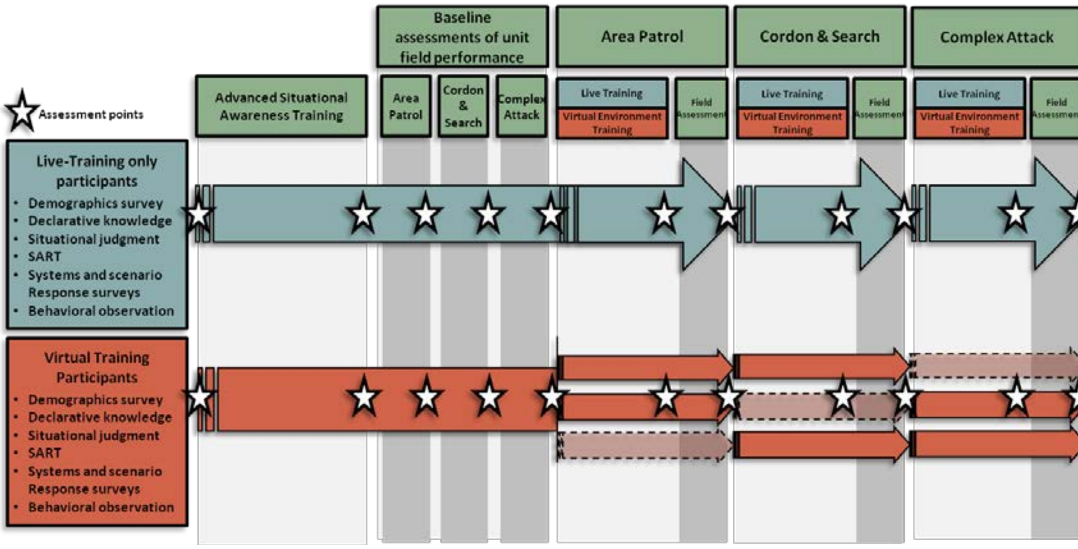

Figure 1. Experimental Treatment flow, by participant groups.

Virtual and Live Environments

Following establishment of the baseline level of unit performance, each virtual training squad conducted three days of training for their mission in one of the immersive training systems (Dismounted Soldier Training System (DSTS) or Small Unit Virtual Immersion System (VIRTSIM)). These virtual infantry training capabilities both consisted of a wearable computer, a head mounted display, a simulated weapon, a motion tracking system and simulated audio and radio capabilities. Employment of these capabilities was different as a result of their associated movement techniques. DSTS provided for joystick controlled movement within the virtual environment, and hand and arm signals. VIRTSIM allowed for naturalistic movement within a certain volume of space with the trainee pivoting to continue forward within the environment (see Figure 2). VBS2 was the simulation engine that drove the virtual events on a moderate fidelity three dimensional model of McKenna MOUT site. The control squad prepared for each mission by conducting standard live training methods, using the terrain where the record run mission sets were executed (see Figure 3). Like the immersive environment trainees, the control squad received support from role players and the MOUT site's after action review (AAR) capability.

Figure 2. Experimental squad soldiers in the VIRTSIM system conducting an attack on a virtual McKenna MOUT site.

Figure 3. Control squad soldiers conducting an attack on McKenna MOUT site. (Courtesy of Maj. Thanh, USA.)

Collection

When possible, trainee burden was reduced by providing the trainees with a single interface through which to receive the daily tests and surveys: Netherlands Aerospace Laboratory's (NLR) Bold Quest survey tool. Soldiers were provided a user number and password, which they utilized throughout the event to access the survey tool, answer surveys and access tests. Throughout the duration of the experiment, the survey tool was installed on 100+ machines, and all were leveraged for collection regularly; over 66,000 data points were collected.

A total of 220 hours of squad communications were captured during MOUT and Virtual training, with accompanying Position Location Information (PLI) and virtual PLI-equivalent data. Ninety hours of that data included key mission performances captured during baseline week, and the record runs performed by each squad in live as well as virtual environments. The volume of data collected helped to make up for some of the limitations that are inherent in data capture efforts associated with a large experiment in a synthetic operational training environment. These 'limitations' proved to be multipliers in some circumstances. All Soldiers and Marines were given a breadth of exposure across multiple systems, rather than a deep exposure just to one system. While the system rotation created its own non-statistically significant confounds, it also exposed each capability to a larger audience with diverse warfighter experiences.

PARTICIPANT DEMOGRAPHICS

Table 1. Demographics of BQ12-2 Participants

Participants	Avg. Age	Avg. Yrs. in Service	Avg. No. of Deployments	Combat Months
USA Control Squad (n=9)	22.22	3.00	.67	9.22
USA Experimental Squad (n=10)	25.27	5.50	1.18	15.70
USMC Squad (n=14)	22.71	3.67	1.00	12.43
Canadian Army Section (n=11)	28.64	6.82	.64	6.23

The initial demographics collection included 19 U.S. Army (USA) soldiers, 14 U.S. Marine Corps (USMC) Marines, and 11 Canadian Army soldiers (see table 1). Participant numbers varied throughout the experiment as a factor of the participation level of each squad's leadership, and the availability of 'spare' unit members. Where appropriate, unit leadership and soldiers who were not part of the weekly missions still provided feedback on the training they had experienced. The Canadian Army Section participating was an organic unit while the USA squads and USMC squads sourced their soldiers from within a company. These soldiers were familiar with each other and had shared Standard Operating Procedures (SOPs) and TTPs, but had not necessarily deployed together.

Other items were collected to analyze the impact of personal experiences on the training. This included soldiers' previous training in areas that might confound the collection, experience with computers and training simulations, as well as hobby-level participation in both computer oriented and outdoor oriented activity. Thirty-seven of the forty-four participating soldiers stated that they played video games at least once a week. A total of twenty-four specifically cited that they had most recently played either Battlefield 3, or Call of Duty: Modern Warfare.

Another area explored was trainee perceived utility of particular types of training, as well as their familiarity and previous exposure to that training type. On the scale provided to trainees, 1 was rated as "extremely not relevant"; 2 was "not relevant"; 3 was "somewhat not relevant"; 4 was "neutral"; 5 was "somewhat relevant"; 6 was "relevant"; and 7 was "extremely relevant". Scenario-based training was viewed more highly by the USA control squad (μ=5.22) and the USMC squad (μ=4.93), with both scores close to "somewhat relevant", than by the USA Experimental (μ=3.72) Squad and the Canadian Army section (μ=4.12). The difference was not statistically significant.

A total of 48 soldiers said they had no hands-on familiarity with virtual training concepts; of those that were familiar with virtual training, they rated the relevance of the training as neutral, with a combined mean of 4.1. Unfamiliarity with formal military virtual training did not imply unfamiliarity with commercial virtual systems. This stated familiarity with highly polished commercial titles which simulate some level of military operational fidelity may have factored into trainee expectations, even if it did not reveal itself as a statistically significant covariate.

SITUATIONAL AWARENESS AND SITUATIONAL UNDERSTANDING

Table 2. SART Scales and Questions (Taylor, 1989).

Supply of Attention Resources	Demand on Attention Resources	Understanding of the Situation
• How **aroused** are you in the situation? Are you alert and ready for activity (high) or do you have a low degree of alertness (low)? • How much are you **concentrating** on the situation? Are you bringing all your thoughts to bear (high) or is your attention elsewhere (low)? • How much is your **attention divided** in the situation? Are you concentrating on many aspects of the situation (high) or focused on only one (low)? • How much **mental capacity** do you have to **spare** in the situation? Do you have sufficient to attend to many variables (high) or nothing to spare at all (low)?	• How **changeable** is the situation? Is the situation highly unstable and likely to change suddenly (high), or is it very stable and straight forward (low)? • How **complicated** is the situation? Is it complex with many interrelated components (high) or is it simple and straightforward (low)? • How many **variables are changing** in the situation? Are there a large number of factors varying (high) or are there very few variables changing (low)?	• How much **information have you gained** about the situation? Have you received and understood a great deal of knowledge (high) or very little (low)? • How **good is the information** you have gained about the situation? Is the knowledge communicated very useful (high) or is it a new situation (low)? • How **familiar** are you with the situation? Do you have a great deal of relevant experience (high) or is it a new situation (low)?

Units were assessed on the demands placed on unit and individual attention resources; changes to individual supply of attention, when presented with similar situations across different training capabilities; and the transition of awareness to understanding. This awareness then transitioned, or did not transition, to a unit's shared understanding. Concepts investigated included whether individuals were able to accurately identify which information was relevant, and then whether the unit leader was able to apply analysis and judgment to information deemed relevant. Also collected was the state of common squad understanding of the mission, as well as shared expectations of the tasks, and understanding of their own roles and responsibilities.

The SART is a 10 item Likert-type scale that allows the trainee to rate their self-assessment of the situation they just completed, and their ability to function within that scenario (Taylor, 1990). These ten items are then broken down into three scales, which encompass the trainee's self-rating of the event's demand on their attention resources, their supply of resources during the event, and their understanding of the event (Table 2). Observational data supported the situational awareness rating technique data, user feedback in the forms of surveys, hot washes and interviews, as well as PLI and other performance measurements.

Results

When utilizing the SART, SA is calculated as demand on attention minus supply of attention, subtracted from the score of the understanding dimension. Trainees were administered SART after their main effort mission run in the live scenarios, and after their record mission run on the last training day of each week. Completion of this survey was encouraged but not mandatory, though overall compliance was high. Due to the exercise rotation, which caused each squad participating in the experimental virtual training treatment to be unavailable for one week's final assessment on a rotating basis, the three squads that composed the virtual training treatment group were not analyzed as a single entity (table 3). When using a repeated measures ANOVA, the mean score for within-subjects effects of SA is statistically significant, $F(3, 960)= 281.658$, $p=.000$, $\eta p^2= .468$. The within subjects interaction of the four dimensions of situational awareness and mission was also statistically significant, $F(9, 960)=3.875$, $p=.000$, $\eta p^2= .057$. This correlated to the observational data collected throughout the experiment. While the scenarios during the post-training assessments were more challenging than the baseline assessments, participant self-assessment of situational awareness declined to a lesser degree than expected due to the increased challenge. Towards the end of the missions, trainees stated they began more purposefully applying their understanding of the environment onto their next action, though the actual execution was not always effective.

Between subjects, the dimensions of understanding, supply of attention and overall situation awareness were different to a statistically significant degree for the squads (Understanding: $F(3, 428)= 3.105$, $p=.026$, $\eta p^2=.021$;

Supply: $F(3, 428)= 3.914$, $p=.009$, $\eta p^2 = .027$; Situational Awareness: $F(3, 428)=4.357$, $p=.005$, $\eta p^2=.03$). This statistically significant difference crossed the bounds between the experimental and control groups, as both the USA control squad and the Canadian Army Section's reported understanding, supply of attention and overall SA were higher than the USA experimental squad's on a mission-to-mission basis. For overall SA, the USMC squad was also a statistically significant scorer as compared to the USA experimental squad, with similar levels of understanding but lower levels of attentional supply than the USA experimental squad. Demand on attention did not differ to a statistically significant degree between squads; it did differ as a factor of mission, $F(3, 428)=2.866$, $p=.036$. Trainee SA dropped, due to the increased challenges associated with the final missions. The type of training capability that soldiers were participating in weekly did not have a statistically significant effect on the outcome of their SA, while SA changes correlated to mission and time.

Table 3. Calculated Soldier SA Across Training Events

		Understanding	Demand	Supply	SA
ASAT Training	**USA Control Squad (n=9)**	14.76	12.06	17.59	20.29
	USA Experimental Squad (n=11)	14.20	13.45	17.65	18.40
	Canadian Army Section (n=11)	17.40	15.40	19.95	21.95
	USMC Squad (n=14)	14.76	14.60	17.44	17.60
Baseline Area Recon	**USA Control Squad (n=9)**	14.11	10.89	20.78	24.00
	USA Experimental Squad (n=11)	14.29	12.29	20.71	22.71
	Canadian Army Section (n=11)	13.67	12.67	20.89	21.89
	USMC Squad (n=14)	14.29	10.79	18.79	22.29
Baseline Attack	**USA Control Squad (n=9)**	16.25	13.00	22.38	25.63
	USA Experimental Squad (n=11)	14.75	14.38	19.38	19.75
	Canadian Army Section (n=11)	13.67	14.33	20.89	20.22
	USMC Squad (n=14)	14.23	14.62	19.46	19.08
Area Recon	**USA Control Squad (n=9)**	15.00	13.28	19.44	21.17
	USA Experimental Squad (n=11)	12.42	10.89	15.16	16.68
	Canadian Army Section (n=11)	13.60	15.15	18.90	17.35
Cordon & Search	**USA Control Squad (n=9)**	14.39	14.11	20.33	20.61
	Canadian Army Section (n=11)	13.10	12.40	17.60	18.30
	USMC Squad (n=14)	15.46	13.88	18.31	19.88
Attack	**USA Control Squad (n=9)**	13.82	13.41	18.94	19.35
	USA Experimental Squad (n=11)	14.12	12.24	18.12	20.00
	USMC Squad (n=14)	12.81	12.56	15.04	15.30

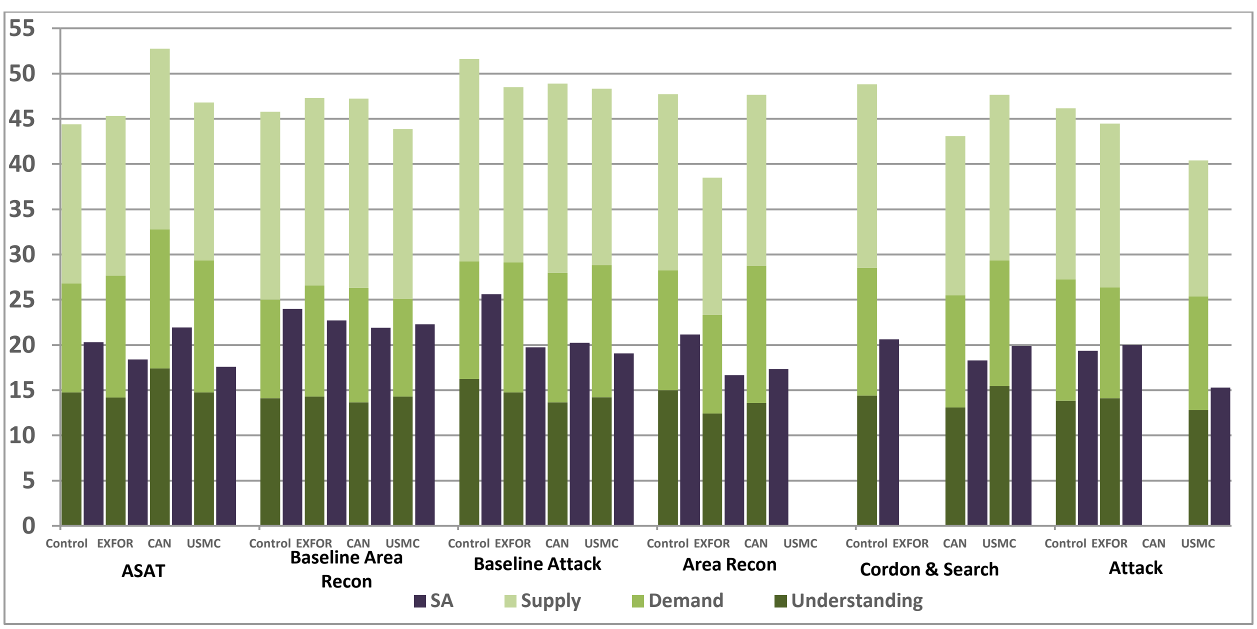

Figure 4. Soldier SART self-rates, by Activity.

Outcomes

Differences in individual situational awareness were not impacted by the specific training capability each squad utilized. Rather, changes were impacted by the mission executed in a given week. Changes were also brought about by the tactical level performance of each unit, though this correlation was only reaching towards statistical significance, not achieving it. While a weakness of utilizing the SART methodology is the subjective nature of self-report data, results did not correlate to the self-efficacy self-rate scale administered at the same time. Neither did the self-efficacy self-rate scale data correlate to trainee mission performance, which could suggest that the SART scenario-responsive reports were not impacted by social desirability factors.

In addition to the situational awareness dimension scores, data was collected on shared role perception, as well as mission understanding. In both the experimental and control groups, roughly one trainee per event, was unable to accurately identify the mission being conducted. Squad leaders more frequently provided detailed information on the mission just completed, as well as the main tasks and scenario role. Lower level enlisted and those with lower levels of personal experience focused on their individual tasks and individual performance levels throughout the event.

Soldier SA changed not as a factor of the system they were in but as a factor of the mission they were performing. The lack of variance between squads emphasized the evenness of impact the training capabilities had on trainee SA – whether that training was live or one of the two primary virtual systems.

Inter- and Intra-Squad communications

An integral component of situational awareness and team performance is the communications of the small unit and individuals within the small unit. Measures of merit explored included the frequency of the squad leader pulling information from his unit; instances of communications being pushed; quality of the communications themselves; whether the information loop was closed appropriately; and if information was acted on appropriately (Wilson, et al, 2007).

Communications were sourced from a tape recorder placed with the squad leader during the main effort and virtual recorded mission completions during the baseline week and each training week. This collection was impacted when the squad leader "died" in scenario. Each instance was noted as a single item if it occurred within the same 60 second period (i.e., a squad member being contacted by the squad leader but not responding to three attempted contacts is counted as one if it was part of the same continuous incident).

Results

Changes to communications outcomes were statistically significant within subjects (table 4), $F(4,36)=66.76$, $p=.000$, $\eta p2= .881$. Within-subjects effects of the interaction of communications outcomes and treatment was also significant, $F(12, 36)= 2.099$, $p=.043$, $\eta p2=.412$. There were no statistically significant differences between-subjects for squads, treatment, or event. Breaking down the dimensions of communications outcomes, for a one-way ANOVA, the interaction of treatment and communications was significant for communications pushed to leader $F(3, 29)=4.832$, $p=.008$, as well as the loop of communication being closed, $F(3,19)=4.957$, $p=.007$.

Table 4. Squad Communications Changes, by Event. (T*: Training event; P*: Post-training event)

Communications Pulled by Squad leader									
Squad	Area Recon	T Area Recon	P Area Recon	Cordon&Search	T Cordon&Search	P Cordon&Search	Attack	T Attack	P Attack
US Army Experimental Squad (n=9)	2	5	2	3	-	-	2*	6	2
US Army Control Squad (n=9)	3	4	2	2	3*	2	6	3	3
Canadian Army Section (n=11)	3	9	4	5	4	1*	2	-	-
USMC Squad (n=14)	2			7	8	0*	4	13	4
Communications Pushed to Squad leader									
Squad	Area Recon	T Area Recon	P Area Recon	Cordon&Search	T Cordon&Search	P Cordon&Search	Attack	T Attack	P Attack
US Army Experimental Squad (n=9)	4	8	10	7	-	-	6*	12	14
US Army Control Squad (n=9)	7	13	8	7	5*	22	8	16	18
Canadian Army Section (n=11)	9	14	13	4	8	10*	3	-	-
USMC Squad (n=14)	6	-	-	8	22	8*	9	6	10

Communications pushed, unacknowledged									
Squad	Area Recon	T Area Recon	P Area Recon	Cordon&Search	T Cordon&Search	P Cordon&Search	Attack	T Attack	P Attack
US Army Experimental Squad (n=9)	1	0	1	0	-	-	2*	0	1
US Army Control Squad (n=9)	0	2	0	1	0*	1	2	0	2
Canadian Army Section (n=11)	2	6	2	2	3	2*	2	-	-
USMC Squad (n=14)	0	-	-	2	2	0*	1	2	1
Garbled Communications									
Squad	Area Recon	T Area Recon	P Area Recon	Cordon&Search	T Cordon&Search	P Cordon&Search	Attack	T Attack	P Attack
US Army Experimental Squad (n=9)	0	0	0	0	-	-	2*	0	0
US Army Control Squad (n=9)	1	0	0	1	0*	1	0	0	3
Canadian Army Section (n=11)	1	0	0	0	0	0*	2	-	-
USMC Squad (n=14)	1	-	-	4	0	0*	3	0	0
Loop closed on communications									
Squad	Area Recon	T Area Recon	P Area Recon	Cordon&Search	T Cordon&Search	P Cordon&Search	Attack	T Attack	P Attack
US Army Experimental Squad (n=9)	5	12	11	3	-	-	4*	18	15
US Army Control Squad (n=9)	10	15	10	8	8*	23	12	19	19
Canadian Army Section (n=11)	10	17	15	7	9	9*	3	-	-
USMC Squad (n=14)	8	-	-	13	28	8*	12	17	13

*Squad Leader Died During Mission Execution

Outcomes

Instances of communications being pushed to the squad leader increased after training treatment – regardless of whether the training treatment was live or virtual training in either of the two immersive systems. Within the virtual systems, squad leaders performed more pulling of communications, but in live scenario performance, their pulling of communications dropped to the same level as that observed with the control squad. Trainees self-reported surveys that they spent more time communicating in the virtual environment as an attempt to make up for the subtle visual cue losses associated with executing their mission in a virtual environment. All participants experienced an increase in clarity of communication, and in items of information being passed to the squad leader. For improving communication consistency within the squad, the virtual systems ultimately yielded similar gains as the live training.

FUTURE RESEARCH

The results of the data collected at AEWE-BQ12-2 revealed gaps in the understanding of how virtual and immersive systems interact with each-other and impact the Warfighter both in comparison to live training and as a supplement to live training in a blended learning process. The literature on immersive and virtual training as it stands today represents a body of comprehensive analytical work, to which the results of AEWE-BQ12-2 adds yet more questions. While sections of the scientific community of interest work to answer these questions, there has been no unified effort led by the DOD to understand the trade-offs in task fidelity provided by the larger body of virtual capabilities across DOD and industry, as well as the level of fidelity required to train tasks to proficiency.

As part of the lessons learned during AEWE-BQ12-2, the following areas for further research were suggested:

- The utility of desktop virtual training contrasted to immersive virtual training systems for the dismounted squad. While this work has been explored in Taylor and Barnett (2012) on a naive, civilian training audience with demographics comparable to basic trainees, a similar tightly controlled experiment has not been performed on experienced warfighters drawn from appropriate Military Operational Specialties.
- An assessment of the optimal mixes of Live, Virtual, and Immersive Virtual training (as well as constructive and gaming) by task types.

CONCLUSIONS

Current training environments impact small unit SA, SU, and readiness best when they create a learning environment that the squad leader is comfortable manipulating. This was realized in the AAR capability of VIRTSIM and DSTS, where soldiers were able to experience the ground truth of their performance in scenarios as contrasted to their perceived performance. The ability to model performance and perform quick repetition outweighed the occasional lack of a convincing survival threat. For changes which relied on visual discrimination, the two immersive virtual systems posed a challenge. Many times throughout the experiment, changes in performance were less traceable than changes in behaviors and group norms which would, given time, eventually lead to improvements in the desired performance.

The changes in squad communications experienced in the virtual environments matched those experienced in live training. This concurred with soldier feedback stating they were able to work on and improve communications in the virtual environment, despite or perhaps because of the challenges posed by the environment-supplied communication systems.

Taken individually, almost every system soldiers and marines trained with had an impact on some measure of soldier awareness, readiness, communications, and friendly, enemy, neutral and non-combatant sorting. While it did not translate into statistically significant gains in mission performance over the 5-week experimental period soldiers who went through the immersive training were able to maintain their performance in increasingly challenging live missions at a level on par with those who experienced live training. This indicates a trend which could be realized if these capabilities were inculcated U.S. Army wide.

In today's acutely resource constrained environment – both fiscal and in terms of encroachment on key training ranges – it is necessary to analyze the gains afforded by utilization of virtual systems on combat readiness. As these systems continue to improve in visual fidelity and the ability to accurately model human behavior in the operating environment, virtual training systems will find their place in the comprehensive training continuum. It is the DOD's responsibility to make sure that the capabilities provided to conduct and enhance small unit training are sufficient to train in an operationally realistic manner, provide adequate throughput, and allow small unit leaders to develop their unit performance in a manner that is responsive to its needs. The capabilities observed during AEWE-BQ12-2 reflect a positive first step down that road, but there is considerably more work to be done.

ACKNOWLEDGEMENTS

This work was supported in part by the U.S. Joint Staff (Contract #W91247-11-D-0010). 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 US Joint Staff or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes.

REFERENCES

AUSA Torchbearer. (2011). *National Security Report: The US Army Squad: Foundation of the Decisive Force.* Arlington, VA: Institute of Land Warfare, Association of the United States Army.

De Leo, G., Sechrist, S., Radici, E., & Mastaglio, T.W. (2010). Games for Team Training. *Proceedings of the 2010 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC).*

Department of Defense. (2010). *The Strategic Plan for the Next Generation of Training for the Department of Defense.* Washington, D.C.: Office of the Under Secretary of Defense (Personnel & Readiness).

Department of the Army. (2001). *FM 3-0: Operations.* Washington DC: Department of the Army.

Department of the Army. (2011). *Integrated Concept Development Team (ICDT) Squad (SQD) Capabilities-Based Assessment (CBA).* Fort Benning, GA: United States Army Maneuver Center of Excellence (MCoE).

Dorsey, D., Russell, S., & White, S. (2009). Identical elements theory: Extensions and implications for training and transfer. In J. Cohn, D. Nicholson, & D. Schmorrow (Eds.), *The PSI handbook of virtual environments for training and education: Vol 3: Integrated systems, training evaluations and future directions* (pp. 196–205). Westport, CT: Praeger Security International.

Endsley, M.R. (1995a). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, 37, 32-64.

Endsley, M.R. (1995b). Measurement of Situation Awareness in Dynamic Systems. *Human Factors*, 37, 65-84.

Greitzer, F. L. & Andrews, D.H. (2010). Training Strategies to Mitigate Expectancy-Induced Response Bias in Combat identification: A Research Agenda. In D. H. Andrews, R. P. Herz, & M. B. Wolf (Eds.), *Human Factors in Combat identification* (pp. 173-189). Burlington, VT: Ashgate Publishing Company.

Knerr, B. W. (2007). *Immersive simulation training for the dismounted Soldier. Army Research Institute study report 2007-01*. Arlington, VA: US Army Research Institute for the Behavioral and Social Sciences.

Knerr, B. W., Garrity, P. J., & Lampton, D. R. (2005). Embedded training for future force warriors: An assessment of wearable virtual simulators. *Paper presented at the 24th Annual Army Science Conference*, Orlando, FL

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

Reitz, E. A, Seavey, K. P. (2012). Using Virtual Environments to Improve Real World Performance in Combat Identification. *In Proceedings of the Interservice/Industry Training, Simulation and Education Conference 2012*. I/ITSEC, Orlando, FL.

Roman, P. & Brown, D. (2009). Games – just how serious are they? *Proceedings of the 2009 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC).*

Shrader, C.R. (1982). *Amicicide: The problem of friendly fire in modern war*. Fort Leavenworth, KS: U. S. Army Command and General Staff College.

Spiker, V. A. & Johnston, J. H. (2010). Using behavioral science principles to train small units. *Proceedings of the First Cross-Cultural Decision- Making Conference (CD-ROM).* Miami, OH.

Taylor, G.S. & Barnett, J.S. (2012)**.** Evaluation of Wearable Simulation Interface for Military Training. *Human Factors: The Journal of the Human Factors and Ergonomics Society* published online 29 Nov. 2012.

Taylor, R.M. (1990). "Situational Awareness Rating Technique (SART): The Development of a Tool for Aircrew Systems Design". In *Situational Awareness in Aerospace Operations*, AGARD-CP-478. 3/1 - 3/17. Seuilly-sur-Seine, France: NATO-AGARD.

US Army Training and Doctrine Command. (2010). *The United States Army Learning Concept for 2015 (TRADOC PAM 525-8-2).* Fort Monroe, VA: Department of the Army, Training and Doctrine Command.

US Army Training and Doctrine Command. (2011). *The United States Army Training Concept 2012-2020 (TRADOC PAM 525-8-3).* Fort Monroe, VA: Department of the Army, Training and Doctrine Command.

US Congress, Office of Technology Assessment (OTA). (1993). *Who Goes There: Friend or Foe?* Washington, DC: US Government Printing Office.

US Joint Forces Command. (2011). *Joint Cooperative Target Identification - Ground (JCTI-G) Analysis of Alternatives (AoA).* Norfolk, VA: US Joint Forces Command, JS C4 Joint Fires Division.

US Joint Forces Command. (2010). *Future Immersive Training Environment Joint Capability Technology Demonstration Spiral 1 Limited Joint Operational Utility Assessment Report. Suffolk, VA:* US Joint Forces Command, J7.

Wilson, K.A., Salas, E., Priest, H.A., Andrews, D. (2007). Errors in the Heat of Battle: Taking a Closer Look at Shared Cognition Breakdowns Through Teamwork**.** *Human Factors: The Journal of the Human Factors and Ergonomics Society*, April 2007, 49: 243-256.