# Proficiency Evaluation and Cost-Avoidance Proof of Concept M1A1 Study Results

**Robb Dunne**
**Innovative Reasoning LLC**
**Orlando, FL**
**Robert.Dunn@InnovativeReasoning.com**

**Dr. Tim Cooley**
**DynamX Consulting**
**Larkspur, CO.**
**TCooley@DynamX.org**

**Dr. Steven Gordon**
**Georgia Tech Research Institute**
**Orlando, FL**
**Steve.Gordon@gtri.gatech.edu**

**ABSTRACT**

Evaluation of simulation-based training systems to determine their contributions to trainee proficiency and to determine the level of cost avoidance vice live training is essential to plan the future live-virtual-constructive training environment for the United States Marine Corps (USMC). This need is reinforced in a recent Government Accountability Office report (GAO 13-698, August 2013) on Army and Marine Corps Training titled *Better Performance and Cost Data Needed to More Fully Assess Simulation-Based Efforts*, which states that the Services "lack key performance and cost information that would enhance their ability to determine the optimal mix of training and prioritize related investments." USMC Program Manager Training Systems (PM TRASYS) has conducted cost avoidance studies on USMC simulation-based training systems for the past 2 years, and these studies are being refined to capture improved cost information. A related study, begun in June 2013, evaluates the effects of USMC simulation-based training programs on proficiency.

This paper presents the process, results, and recommendations of the recent PM TRASYS Proof of Concept (POC) study of measuring proficiency changes and cost avoidance due to use of the M1A1 Advanced Gunnery Training System (AGTS) simulator. For the POC, a group of consistent crews in initial AGTS training are monitored (without interference) through a sequence of 10 gunnery table tasks, with a total of 500+ task instances in the AGTS simulator, to the culminating live-fire tasks. Early session scores are compared to "Gate-To-Live-Fire" scores in the simulator, and these results are compared to the live-fire M1A1 qualification scores for these crews. Results of the POC are promising. The study finds that with performance-oriented metrics and measures, tied to doctrine and captured automatically, it is possible to determine both proficiency trending and cost avoidance. This paper also discusses lessons learned and provides recommendations and implications of findings for training system design.

**ABOUT THE AUTHORS**

**Mr. Robb Dunne** is the Instructional Systems Specialist for the Combat Vehicle Training Systems Program under Program Manager Training Systems, Marine Corps Systems Command and is completing his dissertation: Empirically Defining Scenario Complexity: Toward Adaptive Automated Scenario Generation for his PhD in Instructional Technology from University of Central Florida.

**Tim Cooley, PhD** is President and Founder of DynamX Consulting, a veteran owned consulting firm located in Larkspur, Colorado. Dr. Cooley spent 16 years on the United States Air Force Academy faculty, both in uniform and as a civilian, holding numerous positions including DMSO Modeling and Simulation Chair, Deputy Department Head, and Senior Researcher.

**Steve Gordon, PhD** is the Orlando Field Office Manager for Georgia Tech Research Institute and the Director of Georgia Tech's Test and Evaluation Research and Education Center. He served 26 years in uniform in the United States Air Force and 5 more years as a federal government civilian employee as Technical Director for the Air Force Agency for Modeling and Simulation.

# Proficiency Evaluation and Cost-Avoidance Proof of Concept M1A1 Study Results

**Robb Dunne**
**Innovative Reasoning LLC**
**Orlando, FL**
**Robert.Dunn@InnovativeReasoning.com**

**Dr. Tim Cooley**
**DynamX Consulting**
**Larkspur, CO.**
**TCooley@DynamX.org**

**Dr. Steven Gordon**
**Georgia Tech Research Institute**
**Orlando, FL**
**Steve.Gordon@gtri.gatech.edu**

## DEMAND SIGNAL

The Army and Marine Corps use live and simulation-based training to meet training goals and objectives. Officials for both Services have noted benefits from the use of simulation-based training, in terms of training effectiveness and in terms of cost savings or cost avoidance (Government Accountability Office, 2013). Available findings show that simulators are cost-effective and provide training as beneficial as the use of the actual equipment (Bennell & Jones, 2004).

However, while the Army and Marine Corps collect data on the usage of simulators to measure economic impact of simulator use, a recent Government Accountability Office (GAO) report states that neither Service has established metrics or indicators to assist them in measuring the impact of simulation-based training on improving the proficiency of warfighters. The GAO report also states that, "Without a means to assess the impact of using simulators on performance and to compare the costs associated with live training and the use of simulation-based training devices, decision makers in the Army and Marine Corps lack information to make fully informed decisions in the future regarding the optimal mix of training and related investment decisions" (GAO, 2013).

Traditionally, these questions have been answered by Training Effectiveness Evaluations (TEE) and Cost Analyses of training systems. Unfortunately, a TEE does not typically address proficiency; only whether or not the system is effective at delivering training which allows the trainee to perform at or above standard. This degree of effectiveness is often defined in a binary fashion. That is, a TEE determines the *end state* -whether the trainees do or do not pass some criteria after training in the system. The Department of Defense (DoD) defines proficiency as, "the ability to perform a specific behavior, task or learning objective to the established performance standard in order to demonstrate mastery of the behavior of crews" (DoD, 2001). Thus, while it can be determined that performance standards are achieved, a TEE does not determine or illustrate the impact of the simulator upon trainee proficiency *during* the progress of training. This is a subtle but vital difference.

Additionally, traditional Cost Analyses have not been used in conjunction with analyses of proficiency. That is, Cost Analyses do not currently have the capability to answer the question; how much cost was avoided during the process of bringing trainees from non-proficient to proficient or combat ready?

## OVERVIEW

This paper reports on the results of a Proof of Concept (POC) conducted by Program Manager Training System (PM TRASYS), a component of Marine Corps Systems Command (MARCORSYSCOM), to address recommendations contained within GAO report 13-698, *Better Performance and Cost Data Needed to More Fully Assess Simulation-Based Efforts* (GAO, 2013). The POC assesses the impact of simulator-based training on the proficiency of crews in the Abrams Main Battle Tank (M1A1) Advanced Gunnery Training System (AGTS). The POC explores the feasibility of using simulation-based training scores for correlation to trainee proficiency as measured in live-fire M1A1 crew qualification scores. Demonstration of increased trainee proficiency is combined with existing cost avoidance calculation methodology to illustrate the value of this training system.

The POC satisfies the GAO recommendations to: 1) identify trainee performance-oriented metrics in order to assess the impact of simulation-based training on improving the proficiency of service members and units and 2) develop a methodology to measure cost avoidance of use of simulation-based training vice live training. Additionally, the POC confirms a methodology previously used that allows identification of costs avoided by using simulator training to increase proficiency.

The POC leverages doctrinal, well-defined metrics for M1A1 Gunnery Table (GT) VI tasks, that form the basis of M1A1 Live-Fire Qualification (LFQ), and associated measures captured by the AGTS instructional software. The United States Marine Corps (USMC) Tank community provided vital cooperation with data collection, participation of trainees, and use of the M1A1 AGTS simulator that serves as the test-bed for the POC.

In order to achieve the objectives of the POC, three sources of data are required: 1) AGTS performance data, 2) simulator usage data, and 3) M1A1 LFQ results. Per USMC doctrine (Department of the Navy, 2013), M1A1 trainees perform a series of AGTS exercises encompassing GT VI tasks and culminating in successful completion of the Gate-To-Live-Fire (GTLF) exercise (GT VI performed in the simulator) prior to LFQ. Performance metrics, listed as scores by tasks and manually captured and printed from the AGTS Instructor Operator Station (IOS), are collected to calculate measures of proficiency. Criteria such as time-to-ID, time-to-kill, and other results per established standards form the basis of these task scores. While AGTS software does not store these scores for archiving or future retrieval, trainer usage data, such as quantity and type of rounds expended and simulated miles driven, is automatically collected and stored, to enable calculation of cost avoidance. To evaluate trainee transfer of proficiency from the simulator to LFQ, trends in scores, correlation of simulator scores to live-fire scores, and qualification passing rates are evaluated. LFQ results are manually scored, recorded at the firing range and stored at Battalion offices.

## METHODOLOGY

The methodology the POC employs follows two principles: First, the operational testing community emphasizes the use of measures of *combat* performance such as engagement or battle outcomes. For simulation training systems, analogous variables include the measure of trainee performance in the simulator in relation to combat objectives, the measure of transfer of training from the simulation to the real world, and the measure of the impact of simulator training on performance in live simulation (Simpson, 1999). Second, often the most valid measure of job proficiency is on-the-job performance in which weapons systems and related equipment are used as they would be used in combat. When not possible, the next best alternative is a virtual simulation (Morrison & Hammon, 2000).

Data is taken from gunnery exercises performed during the period of July 2013 to October 2013. Per the Heavy Brigade Combat Team (HBCT) scoring matrices, performance data collected includes target type (armored, light armored, unarmored, and troops), the posture of the M1A1 (offensive or defensive), the range to the target, and the kill time (Department of the Army, 2009). This data is collected manually by the Instructor Operator (IO) who prints out the Performance Analysis, Situation Monitor, and Qualification Performance Analysis reports used for After Action Review (AAR) and containing the data -- but not archived by the AGTS software.

Each M1A1 AGTS GT VI exercise is composed of 10 tasks scored within the range of 0-100. The performance standard is a combined score of at least 700 while passing 7 out of 10 tasks (passing score is 70 in each task) for both the GTLF simulator exercise and LFQ. The exercise receives an overall score that is an aggregate of the scores and associated metrics the AAR reports (described above) contain. Due to the availability of these scores and task-related data, and the fact that task scores are accepted across the tank community as a valid method of measuring ability, it was determined the best way to track proficiency is by using GT VI task-level scoring. For situations where the task score is not automatically calculated, the data collected above is translated into a task score using the HBCT scoring matrices. The AGTS simulator scores are plotted and then compared to their performance in LFQ.

### Proficiency

Proficiency gains and the transfer of this proficiency from the simulation-based training system to combat-related effectiveness are evaluated to form the foundation of claims of cost avoidance, benefits of simulation, and the value of simulation-based training. Training scores from initial and basic, to advanced and GTLF AGTS exercises are tabulated by crew then, using the same metrics and as the touchstone to combat readiness skills, the LFQ scores are also tabulated by crew.  For the POC, initial scores on the tasks are grouped as if they are a “pre-training” GTLF series to compare to the actual GTLF scores.

Each exercise situation is matched to one of the GT VI tasks 60 through 69 (see Table 1) and task scores on each exercise from initial gunnery to GTLF and through LFQ are compared. AGTS and LFQ use similar metrics (i.e. kill time and range) to determine GTLF exercise and LFQ scores (Lockheed Martin, 2012). Kill time is an aggregate score comprised of time-to-target acquisition, time-to-fire, and time for the munitions to fly to the target. Ultimately, performance assessments of GTLF exercises and AGTS GT VI tasks are equivalent to the scoring on GT VI conducted during the LFQ. If the simulator and LFQ scores are positively correlated, then it can be shown that trainee proficiency in the simulator appears to transfer to LFQ (real-world, combat-related skills).

**Table 1. GT VI and GTLF Tasks and Descriptions**

| Task | Description/Notes |
|---|---|
| 60 | Vehicle Commander, a 2XXXXX exercise number or a "0" task on gunnery tables 2 through 5 |
| 61 | Machine Gun Pure with Chemical, Biological, Radiological and Nuclear (CBRN) conditions |
| 62 | Machine Gun Pure no degradations |
| 63 | Main Gun Pure with single target moving or stationary |
| 64 | Main Gun Pure with multiple moving or stationary targets |
| 65 | Change of weapon system with Kinetic Energy (KE) main gun rounds |
| 66 | Change of weapon system with Chemical Energy (CE) main gun rounds |
| 67 | Degraded mode with multiple moving or stationary targets, main gun pure |
| 68 | Degraded mode with moving or stationary single target, main gun pure |
| 69 | Simultaneous fire, combination of machine gun (COAX) and .50 caliber, no main gun, multiple stationary or moving targets |

**Cost Avoidance**

The trainer usage data already captured by AGTS software enables measurement of cost avoidance. Costs for live ammunition and munitions are obtained from Program Manager for Ammunition (PM AMMO), Fiscal Year 2013 (FY13) Standard Unit Price List. Then, using the data from the AGTS Performance Analyses that lists number of rounds "fired" by type for each task, the cost of firing each AGTS exercise is calculated. The rounds "fired" in the simulator were executed to complete a prescribed training program, based upon the Tank Training and Readiness (T&R) manual (Department of the Navy, 2013) therefore there are no extraneously "fired" rounds that would influence reliability of results.

Additional costs factoring into the cost avoidance calculation are contractor operation and maintenance costs for the AGTS. Costs such as electricity and security for both simulator and live range, simulated fuel expended, and live-fire range maintenance and cleanup costs are not included in the cost avoidance analysis. Although these costs could be included for completeness and would increase the total cost avoided, the costs are minor in comparison to the costs avoided by not firing live rounds and thus have minimal impact on the cost avoidance calculation.

There are two major facets of Return on Investment (ROI) calculations: cost and benefit (sometimes called results). Historically, in order to account for the benefits of Modeling and Simulation (M&S), qualitative metrics were assessed (Oswalt et al., 2011), and the task was limited to collecting measures of these qualitative metrics and a limited collection and analysis of quantitative data. Such studies tend to focus primarily on benefits, and are frequently time-consuming, potentially complicated to justify, and less than convincing without quantitative data. Although qualitative statements like, "Simulation allows mistakes and retraining without the risk level of live training" are true, they do not have the evidentiary value of quantitative measures that are desired by decision makers.

Alternatively, complete ROI studies for training systems can be even more time-consuming and require a high standard of documentation between live and simulator/simulation training. Such documentation may include specifics as fuel consumption, field time-to-maintenance ratios and actual vehicle parts not destroyed or damaged because of use of the simulator. Considering the scope of the POC, it was determined studying only the most robust factors of actual costs of the M&S system and the costs avoided if the same task was performed without the M&S

would demonstrate and validate the worth of the M&S system efficiently. This type of approach, however, often overlooks or does not account for all the potential benefits.

For the POC, we apply the quantitative cost avoidance methodology to the AGTS. Costs avoided by not firing rounds on the live range are calculated and costs incurred for operating and maintaining the simulator are obtained. Subtracting the simulator cost data from the cost avoidance of not firing live rounds gives the estimated net cost avoidance.

In equation form, where $LR_a$ = costs avoided by not firing live rounds, $C_{om}$= simulator operation and maintenance costs, and $CA_{net}$ = net costs avoided, the calculation of cost avoidance is:

$$\Sigma\,[LR_a] - \Sigma\,[C_{om}] = CA_{net} \qquad (1)$$

**Limitations and Constraints**

The POC is limited to three intact crews within a few months period of training at one location through a sequence of approximately 500 training events. Looking at trends in scores as trainees progress in the AGTS, and evaluating ways that proficiency improvements can be investigated, is the first step in identifying AGTS system and process improvements to facilitate more complete, powerful studies across multiple sites.

One critical shortcoming common in many training systems is that they do not typically provide the capability to export data in an easily readable format. This ability to export data is critical for trainers, trainees, and analysts because, especially for research purposes, being able to access data at the "shot level" is essential to being able to link shot performance data to a trainee (Chung, et el., 2011). In the case of the AGTS, this capability is important because the collection of the data necessary for conduct of the POC requires significant human capital and "air gap" processes that may add variability and inconsistency.

Variability in the AGTS scoring data may be attributable to several sources such as human error, measurement systems, or the variations in exercise sequencing and content. The sequence of target types and other conditions is randomized by design within the AGTS software. In addition, during conduct of training, system/device faults or errors such as loss of gunner sighting systems or incorrect operator action may have influence on performance. System/device status records are not captured. If such occurrences are captured, it would be possible to correlate them to score anomalies. Some of the variability in scores could also be due to turbulence in crews and crew coordination, instructor inputs, or selected instructional objectives. Knowing the sources of variability, such as the instructor inputs, instructional objectives, and instructor-directed degradations, would be beneficial in the analysis of proficiency.

The AGTS training program leading up to LFQ at this site was put in place by the Battalion's Master Gunner; the AGTS training sequence for this site was not prescribed by the Tank Community's training offices. Therefore, generalizing research findings from the analysis conducted for the POC to community-wide findings for the M1A1 should be considered carefully. This study will be extended to other sites in the next phases.

Limitations and constraints discovered in the POC are captured so that data modifications for the AGTS can be completed prior to the next phases of this study. The proposed solutions to the limitations and constraints on data include updating complete descriptions of the GT VI tasks in the HBCT, providing numerical scores for all AGTS exercises, including all failed exercises, and identifying all crews and trainees for all events in stored training databases while protecting personal data.

Some best practices and advantages of the AGTS training system may also be considered as limitations and constraints. For instance, AGTS exercises are randomized to avoid "gaming the system" and to train agile, adaptable warfighters. While the exercises are basically similar, the order of the tasks and the types of targets and malfunctions vary to better represent reality. This training advantage may add variability to task performance, and it may be a focus of future research. Also, breaks in AGTS training are not consistent across crews; the ready availability of this training system allows supervisors to direct trainees to other tasks for several weeks or more. While we see no correlation between the length of gaps in training and changes in performance, this topic deserves further study. Further, USMC requires crews to train in the AGTS and complete the GTLF qualification before going through the

LFQ; the AGTS is a training and safety requirement. Also, having a control group that goes to LFQ without AGTS training would be of interest, but the feasibility, safety, and cost of this option must be explored. These are only some of the limitations noted in the AGTS POC Phase I study, and modifications will be undertaken to assist the next phase of this proficiency and cost-avoidance study.

**Assumptions**

The following assumptions influence the findings of the POC. First, all rounds shot in the simulator are necessary for training; since the AGTS exercises are instructor-monitored, it is unlikely that the training includes unnecessary rounds fired. However, even if only 25% of the rounds fired are absolutely necessary (and roughly equivalent to what would be fired in live training), the cost avoidance will still be considerable. Second, simulator and live-fire performance data for the three crews is collected in generally identical ways; for follow-on studies, some measurement system analyses on the AGTS and live scoring systems should be conducted. Follow-on studies will evaluate whether the exercise training matrices used across the M1A1 tank community are sufficiently similar to allow comparisons across sites; that is, the sequencing and number of exercises is similar from one training site to another in order to ensure expanded future studies will have consistency. Efforts will be made to standardize exercises and sequencing for follow-on studies.

**Participants**

Due to budgetary and schedule constraints and base operations tempo, only data from three consistent crews was collected during the POC. The three crews were selected as they entered their initial AGTS training in the summer of 2013 because they would likely remain consistent (no personnel changes) during their initial training in the AGTS through GTLF and LFQ. These three Tank Commander (TC) and Gunner (GNR) crews are part of an active duty Tank Battalion during the period from initial AGTS training to GTLF AGTS training, and through LFQ.

These three TC and GNR crews range in rank from Lance Corporal to Lieutenant and average 25 years of age. Excluding the Lieutenant's 15 years of Service experience, the average length in military service for the remaining five trainees is just over four years. These three crews have previous M1A1 experience in other positions and familiarity from schoolhouse training, but at the start of AGTS training, they were new to their positions and to AGTS simulator training at those positions. Five of the six participants had served in Iraq or Afghanistan, or both. Including the time spent in the AGTS prior to the POC, participants average 25.6 hours in AGTS simulation exercises over their career.

## RESULTS

In evaluating proficiency, the three crews participating in the POC have hundreds of scores across the ten tasks in GT VI; however, a sample of three crews is too small to generalize the impact of AGTS simulator training on proficiency for *all* crews. Nonetheless, the POC concludes that the measures used to score performance in the AGTS and in LFQ are consistent and sufficient to judge transfer of proficiency. Identified trends indicate that the AGTS trains tank crews on the skills necessary to become proficient in the simulator and that the skills transfer to LFQ. In light of the conclusion that the AGTS improves proficiency, trainer usage data is used to estimate how much cost is avoided by doing AGTS training versus the same amount of live training. Tables 1 and 2 summarize the proficiency results, and the cost avoidance calculations are summarized in the paragraphs following the tables.

In Figure 1, the period of time in the AGTS simulator is divided into quarters, and average scores for all three crews are grouped into those quarters. In general, the figure shows that AGTS scores improved over time. Each quarter in Figure 1 represents about three weeks in the semi-annual training schedule in the AGTS simulator. From Figure 1, the average of the scores for each crew tend to increase as they gain more time (and training) in the AGTS.

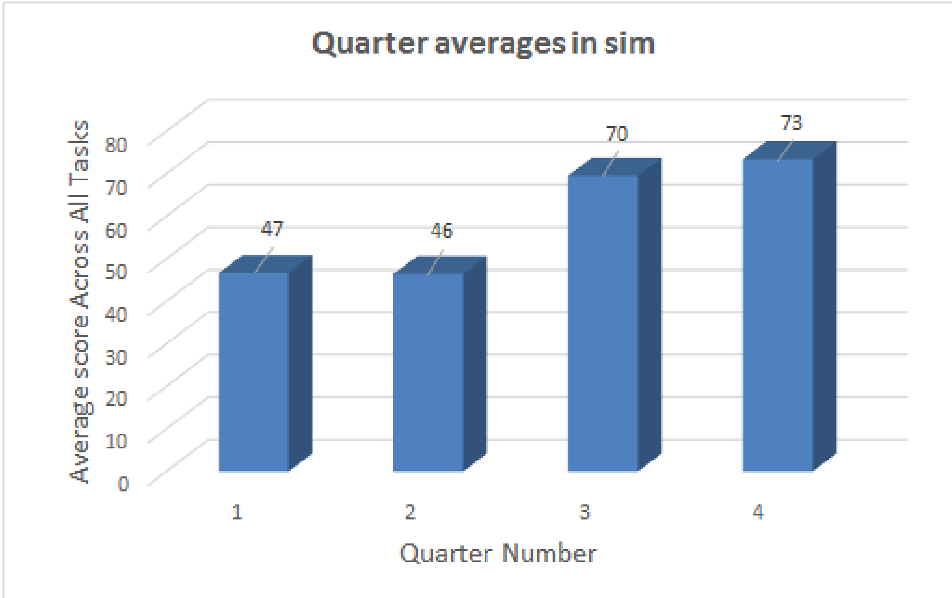


**Figure 1. AGTS Simulator Scores: All Crews, All Tasks**

Results of the POC investigation into if proficiency in the simulator translates to proficiency in LFQ performance are shown in Table 2. Table 2 is a compact view of average crew progress from beginning AGTS simulator scores to ending AGTS simulator scores and then through LFQ scores.

**Table 2. Average Simulator and Live-Fire Scores for All Tasks by Crew**

| Crew | Average Beginning AGTS Scores* | Average Ending AGTS Scores* | Δ = Ending – Beginning Scores | Average Live-Fire Qualification Score* |
|---|---|---|---|---|
| 1 | 63.6 | 93.0 | 29.4<br>(46% Increase) | 90.7<br>(Result: Qualified) |
| 2 | 55.2 | 81.9 | 26.7<br>(48% Increase) | 85.0<br>(Result: Qualified) |
| 3 | 53.6 | 87.0 | 33.4<br>(62% Increase) | 78.0<br>(Result: Qualified) |

***** In conjunction with other evaluation requirements, a passing score is 70.

Depicted in Figure 2 below is an example of trends in Task 67 scores of Crew 1 from their AGTS and LFQ results. Task 67 (as described in Table 1) is chosen here because it is based on degraded modes for AGTS M1A1 systems with multiple moving or stationary targets, and it illustrates scoring trends in one of the more difficult tasks with considerable variability due to system degradation inputs and multiple target insertions. The crew's score is presented on the *y* or vertical axis. The number of exercises is presented along the *x* or horizontal axis. Along with the general trend of increasing proficiency (in conjunction with what appears to be some occasionally challenging mission scenarios), it can be seen that the final AGTS score was quite close to the LFQ score, further demonstrating transfer of skills from the simulator to live-fire performance. There may be several reasons for the near-zero scores such as device malfunction or IO-driven exercise alteration. Without collection of more granular data, it is difficult to identify the cause(s).

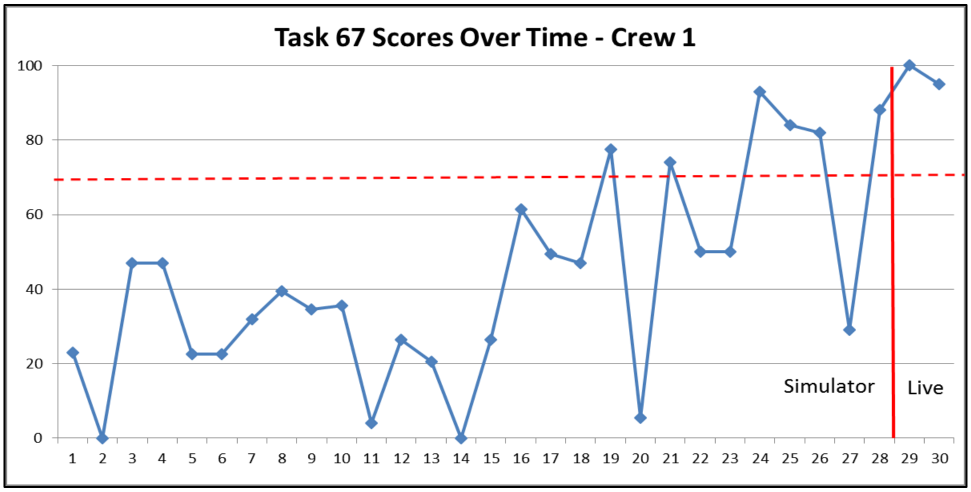


**Figure 2. Task 67 Crew 1 Scores**

On average, all three crews have lower than passing scores on their initial grouping of AGTS training across all tasks. At completion of their GTLF exercises their average AGTS scores show an increase of at least 46%. This improvement in gunnery proficiency could be attributed to the three crews gaining more experience in the AGTS; however, the POC uses LFQ scores as the touchstone to validate proficiency in tasks. In this regard, all three crews pass their LFQ with scores near their average final AGTS scores. Trends in scores by task for each crew indicate that task proficiency, achieved in the AGTS, transfers to LFQ.

Trends in the POC indicate that the task scores provided by the AGTS and from LFQ, are appropriate reflections of performance for use in conducting proficiency studies. Therefore, the objective of the POC is satisfied, and task scoring systems embedded in the AGTS can be used for future proficiency studies. This methodology can also be applied to other USMC training systems that have similar task scoring systems. For the AGTS, an improved, more automated means of collecting proficiency and cost avoidance data is needed to unburden trainers from manual data gathering and to increase the efficiency of data gathering and analysis.

Table 3 below shows the rounds "fired" in the AGTS by crew and the total cost if they were fired as live rounds or live training rounds (whichever is less costly). While tank operations costs are not included here, they are captured in other cost avoidance studies of training systems in the PM TRASYS portfolio.

**Table 3. Number of Simulated Rounds Fired by Crew and Cost Avoided***

| Crew | Munition 1 | Munition 2 | Munition 3 | Munition 4 | Munition 5 | Munition 6 | Totals |
|---|---|---|---|---|---|---|---|
| 1 | 729 | 748 | 7 | 17 | 21 | 110 | 1632 |
| 2 | 460 | 1445 | 4 | 8 | 34 | 75 | 2026 |
| 3 | 999 | 877 | 29 | 36 | 48 | 176 | 2165 |
| Subtotals | 2188 | 3070 | 40 | 61 | 103 | 361 | 5823 |
| Cost | $2,079 | $10,684 | $107,037 | $163,232 | $275,621 | $966,011 | $1,524,663 |

*Munition types omitted for public release

For Cost Avoidance, in FY13 dollars, the cost of the simulator for training these three crews is $7,208, and the costs avoided by "firing" rounds in the simulator totaled $1,524,663, giving a net cost avoidance of $1,517,455 for these three crews over the period of time during the POC. If this training can be assumed to be representative of the training required leading up to qualification, and there are two qualifications per year, then the costs avoided by using the AGTS for training for these three crews per year is $3,034,910.

Extrapolating that yearly amount of cost avoidance to all 116 Active Duty (AD) tank crews in the USMC, the total cost avoided per year by performing training in the AGTS prior to live-fire crew qualification is $117,349,853. Due to time and range availability, LFQ for the 86 Reserve Component (RC) tank crews is conducted only once a year. Cost avoidance per year for RC tank crews performing training in the AGTS prior to LFQ is $43,500,376. Taken together, cost avoidance per year by performing training in the AGTS prior to LFQ across the USMC M1A1 community is $160,850,229. Therefore, in round numbers, it would cost approximately an additional $1M/AD crew, and $.5M/RC crew annually if all qualification training were done on a live range.

It should be noted that this calculation accounts only for the period of the POC, which amounted to approximately three months for one qualification. LFQ is semi-annual for AD and annual for RC. Thus, assuming similar periods of lead-in training, LFQ for AD crews would only account for 6 months and LFQ for RC Crews would only account for 3 months. As usage data gathered from many sites over parts of 2013 and 2014 indicates, trainees use the AGTS year-round, and the above cost avoidance total may reflect only half of the total cost avoidance. In addition, the POC evaluated proficiency and cost avoidance for only the period of time it took these trainees to achieve proficiency in the AGTS and complete LFQ. Trainees with approximately similar experience levels and capabilities as those followed for this POC could require a longer training program if the training was conducted live (not in the AGTS) because of range availability and safety concerns, for instance, potentially increasing live training costs.

## DISCUSSION

In order to readily and consistently provide commanders and other decision-makers with simulation-based data related to proficiency and cost avoidance, the methodology and approach outlined here may benefit other simulation training programs as they seek to answer questions related to proficiency and cost avoidance.

Although cost avoidance data may be readily available from other simulation trainers, trainee performance-oriented measures may be erased after training sessions. The capture of performance-related data may depend on manual extraction and recording. Automated organization, capture, and storage of performance measures of individual trainees and composed crews would enable assessment and forecasting of learning curve, skill, and crew proficiency improvement or decay.

Enhancements to the AGTS software and system and improvements in the AGTS process will be required to track crew progress and log and store scores in a more automated manner. These changes will be made to enable gathering of scores for evaluation of training progress, tracking improvement of performance in the AGTS from the beginning to the end of training, and comparing AGTS GTLF scores to the culminating LFQ scores. Solutions such as Learning Management Systems (LMS), Learning Record Stores (LRS) and statistical analysis software make it possible to track progress with at-a-glance illustrations and automatically store and provide printable output of performance data. On a cautionary note, in order to track student progress for these stated purposes, Personally Identifiable Information (PII) will need to be recorded with scores, and implementation of methods to protect the PII where necessary will be mandatory. Encryption is likely one possible solution.

Automation of these processes would reduce variability and subjectivity of data easily introduced by humans collecting the data and enable objective illustration of trainee skill levels. This information (scores, crew ratings, projected deployment readiness date) could be instrumental in identifying gaps in training, proficiency and deployability. Further, automated processes would supply necessary information upon request within short suspense and with little additional labor dedicated to the task. If standard data presentation formats and graphics are used across the simulator training spectrum, then the need for future government or contracted analytic support could be minimized, and trainers, budgeters, commanders, and other decision makers could use the output as presented.

When developing system requirements such as Key Performance Parameters for a training system, requirements for the system to capture proficiency and cost avoidance metrics and measures in a form ready for analysis should be delineated in the Request for Proposal or Statement of Work, or both. This will ensure the fielding of training systems that enable determination of increases in trainee proficiency and identification of cost avoidances. In this way, during development, the training system will have data collection and delivery capabilities embedded.

In some cases, modifications to existing training systems may improve gathering of proficiency and cost avoidance data in the near term. For example, as result of the POC modifications to the AGTS software are being planned.

Additionally, standardization of an AGTS Program of Instruction (POI) across the M1A1 community may have benefits to training and analysis of proficiency. For instance, conducting the GTLF exercises early in training and then again as the AGTS graduation exercise prior to entering LFQ, would provide standard pre-tests and post-tests for the planning of training and for proficiency analysts to compare scores before and after training. No such standardization, however, is currently in place.

## IMPLICATIONS

The POC looked at a system that is fundamental to training Tank crews. Absent the AGTS, much or all of the same training would have to be conducted live. The use of the AGTS is therefore not only providing a measurable increase in trainee proficiency but also providing significant cost avoidance. The POC also asked questions about the AGTS usage, training processes, crew consistency, crew progression, and trainer data captured. The findings of the POC have implications to how this training system can gather data to help trainers, commanders, and other decision makers understand the impact of the AGTS on trainee proficiency and how training could be modified for individual Marines or crews in the M1A1 and its brother program, the Light Armored Vehicle (LAV) AGTS.

The POC findings indicate that, with the data output made possible through implementation of recommendations in the POC, trainers/instructors may use the improved process to:

- Identify tasks that require more training or less training -- thus determining sources of inefficiencies to improve the overall training system performance.
- Identify crews that need additional or less training; increase task repetitions when crews reach proficiency at a slower rate than planned and decrease task repetitions on tasks where the crew has reached proficiency early.
- Mitigate skill decay by training on identified tasks at an optimal tempo planned to prevent skill atrophy.
- Identify crew compositions that may require adjustment or areas of improvement in crew coordination.
- Incorporate simulator proficiency in the calculation of current unit training level or the timeline to deployment readiness.
- Provide readily available data in an accurate and easy to understand format in order to enable well-founded actions regarding funding, training, and qualification requirements.

The POC results have shown that it is possible to track TC and GNR performance from the gunnery task level in initial gunnery exercises, to GTLF, and then on to live-fire. This type of longitudinal tracking, together with the proper software solutions such as an LMS, LRS or statistical analysis software could enable forecasting of Combat Readiness Percentage (CRP) at differing levels (Platoon, Company, and Battalion).

The POC has shown that the methodology developed can be applied to provide decision makers with an objective assessment of PM TRASYS portfolio training device contributions to warfighter proficiency and training cost avoidance.

The POC also confirms a specified methodology for proficiency estimation using the AGTS. Further studies on other training systems may show the methodology, as modified, could be useful for these purposes across Marine Corps simulator training programs. The ability to judge the effects of simulator training on proficiency is one factor necessary to begin to answer the GAO's recommendations. The other factor is in place by answering the GAO's recommendations on the cost avoidance calculated from using a simulator. As the GAO study recommends, with this proficiency and cost avoidance data, the Marine Corps will improve decision makers' abilities to make fully informed decisions concerning whether training requirements can be met with live and/or simulation-based training and determine optimal mixes of live and simulation-based training in terms of readiness and cost considerations.

## LESSONS LEARNED AND CONCLUSION

The primary lesson learned from the POC effort is that it is possible, with well-defined performance metrics tied to doctrinal field and T&R manuals, to measure proficiency and calculate cost avoidance. The usage data captured automatically on the M1A1 AGTS is critically limited for determination of increases in proficiency. While the usage data provides adequate cost avoidance data, software modifications are needed to collect and store more consistent measures of performance-oriented metrics in order to provide a high-level determination of increases in proficiency. Individual and crew identification within the M1A1 AGTS instructional features is vital for determination of proficiency, readiness, and training plan purposes but must be in keeping with PII protection.

The lessons learned have guided decisions on modifications of the M1A1 AGTS for a more complete follow-on study. This next study will continue to use performance-oriented metrics and measures, tied to doctrine and captured automatically, to determine both proficiency changes and cost avoidance. Also, over the next year, the successful methods used in this study will be used to evaluate other training systems in the PM TRASYS portfolio. For these broader studies, independent review and validation of the results will be appropriate. The results of ongoing cost avoidance studies on many training systems in this portfolio have already influenced requirements stipulated in requests for proposals and contract modifications. As the proficiency studies now expand to beyond the POC for M1A1 AGTS and to additional training systems in the PM TRASYS portfolio, the lessons learned will also influence future requests for proposals and contract modifications. The right data to evaluate proficiency changes and costs avoided should be gathered as automatically as possible, analyzed as needed to influence training delivery and other leadership decisions, and used to help determine training systems' results.

## ACKNOWLEDGEMENTS

The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of PM TRASYS. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon. The authors wish to acknowledge the assistance and cooperation provided by the USMC M1A1 community, Greg Seavers, John Roth and Nathan Jones of PM TRASYS, Larry Miceli of Innovative Reasoning LLC and especially USMC Captain Samuel Oliver for his exemplary leadership and inspiring, relentless dedication.

## REFERENCES

Bennell, C., Jones, N. J. (2004). *The Effectiveness of Use of Force Simulation Training Final Report*. Carleton University. Ottawa, Canada.

Chung, G. K., Nagashima, S. O., Delacruz, G. C., Lee, J. J., Wainess, R., & Baker, E. L. (2011). Review of Rifle Marksmanship Training Research. CRESST Report 783. *National Center for Research on Evaluation, Standards, and Student Testing (CRESST)*.

Department of the Army, Marine Corps Development Command. (2009). Heavy Brigade Combat Team (HBCT) Gunnery FM 3-20.21/ MCWP 3-12.2.

Department of Defense. (2001). MIL-HDBK-29612/4A, Department of Defense Handbook: Glossary for Training. Part 4 of 5.

Department of the Navy. (2013). *Tank Training and Readiness Manual*, NAVMC 3500.121. Washington, D.C.

Simpson, H. (1999). *Evaluating Large-Scale Training Simulations. Volume 1. Reference Manual* (No. DMDC-TR-99-05-VOL-1). DEFENSE MANPOWER DATA CENTER SEASIDE CA DOD CENTER MONTEREY BAY.

Gordon, S., and Cooley, T. (2012). "Phase One Final Report: Cost Avoidance Study of USMC Simulation Training Systems", unpublished report. Georgia Tech Research Institute and DynamX Consulting.

Government Accountability Office. (2013). Army and Marine Corps Training Better Performance and Cost Data Needed to More Fully Assess Simulation-Based Efforts, GAO-13-698.

Lockheed Martin. (2012). *Training System Utilization Handbook, M1A1 Firepower Enhancement Program (FEP) Advanced Gunnery Training System (AGTS) for the USMC Combat Vehicle training System (CVTS).* TM 08953A-10/1. Orlando, FL.

Morrison, J. E., & Hammon, C. (2000). *On measuring the effectiveness of large-scale training simulations* (No. IDA-P-3570). INSTITUTE FOR DEFENSE ANALYSES ALEXANDRIA VA.

Oswalt, I., Cooley, T., Waite, W., Waite, E., Gordon, S., Severinghaus, R., & Lightner, G. (2011). *Calculating return on investment for US Department of Defense modeling and simulation*. DEFENSE ACQUISITION UNIV FT BELVOIR VA.