



ARL-MR-0952 • MAY 2017



Support for Training Effectiveness Assessment and Data Interoperability (STEADI)

**by Ashley Medford, Gabriel Diaz, Jennifer Murphy, and
Gregory Goodwin**

Approved for public release; distribution is unlimited.

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.



Support for Training Effectiveness Assessment and Data Interoperability (STEADI)

by Ashley Medford and Gabriel Diaz
Problem Solutions LLC
Johnstown, PA

Jennifer Murphy
Quantum Improvements LLC
Orlando, FL

Gregory Goodwin
Human Research and Engineering Directorate, ARL

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YYYY) May 2017		2. REPORT TYPE Memorandum Report		3. DATES COVERED (From - To) 05/2015–05/2016	
4. TITLE AND SUBTITLE Support for Training Effectiveness Assessment and Data Interoperability (STEADI)				5a. CONTRACT NUMBER W911NF-14-2-0034	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER 0603015A	
6. AUTHOR(S) Ashley Medford, Gabriel Diaz, Jennifer Murphy, and Gregory Goodwin				5d. PROJECT NUMBER S29	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Problem Solutions, LLC 1407 Eisenhower Blvd., Ste. 303 Johnstown, PA 15904				8. PERFORMING ORGANIZATION REPORT NUMBER ARL-MR-0952	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) US Army RDECOM Army Research Laboratory (ARL) – STTC 12423 Research Parkway, Orlando, FL 32826				10. SPONSOR/MONITOR'S ACRONYM(S) ARMY RDECOM	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT A recent report by the Government Accountability Office found that neither the Army nor the Marine Corps sufficiently assesses the effectiveness of its simulation training systems. To ensure the Army receives the best value for its investment, training effectiveness assessment (TEA) should be a priority for each training system the Army fields. Recent efforts from the US Army Research Laboratory on Interoperable Performance Assessment (IPA) for individuals and teams have made advances on defining and persisting human performance data. IPA has concentrated on leveraging the work on the Experience application programming interface (xAPI) to produce data with intersystem data value. The Support for Training Effectiveness Assessment and Data Interoperability (STEADI) effort conducted the foundational research needed to provide guidance for how to develop effective metrics for TEA, how to represent those measures in xAPI format, how to design and develop an architecture to support the management and persistence of those measures, and an interface so that instructors can easily access the data in meaningful ways. STEADI resulted in the development and preliminary validation of marksmanship measures for TEA, an xAPI Registry with those measures, and the design of an architecture to support their management.					
15. SUBJECT TERMS xAPI, automated assessment, marksmanship, learner record store, Interoperable Performance Assessment					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 34	19a. NAME OF RESPONSIBLE PERSON Gregory Goodwin
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include area code) 407-384-3987

Contents

List of Figures	v
List of Tables	v
1. Introduction	1
1.1 Background	1
1.1.1 Goal	2
1.1.2 Importance	3
2. Research Summary	3
2.1 Interoperable Performance Assessment (IPA)	3
2.2 Human Performance Markup Language (HPML)	3
2.3 Experience API (xAPI)	4
2.4 Soldier Performance Planner (SP ²)	4
2.5 Pipeline	5
2.6 Current Research	5
3. Approach	6
3.1 Project Goals	6
3.2 Project Tasks	6
3.2.1 Task 1: User Needs Analysis	6
3.2.2 Task 2: Measure Development	9
3.3.3 Task 3: Instantiation of xAPI Metrics	11
3.3.4 Task 4: Architecture Design	13
4. Results and Discussion	17
4.1 Overall Findings	17
4.2 Future Recommendations	18
5. Conclusion	20
6. References	22

List of Symbols, Abbreviations, and Acronyms	23
Distribution List	25

List of Figures

Fig. 1	STEADI xAPI registry: main table	13
Fig. 2	STEADI xAPI registry: example xAPI statements	13
Fig. 3	STEADI use case diagram working draft	14
Fig. 4	STEADI Architecture Design Methodology and Tool	15
Fig. 5	STEADI use cases diagram	16
Fig. 6	STEADI architecture diagram	17

List of Tables

Table 1	Rifle marksmanship assessment constructs	10
Table 2	Proposed data collection schedule	11
Table 3	Sample marksmanship measures for Pipeline update	12

INTENTIONALLY LEFT BLANK.

1. Introduction

1.1 Background

As Army training budgets become more limited, a renewed focus has been placed on simulation-based training to maintain Soldier readiness. Through simulation, Soldiers have the opportunity to learn and practice mission critical skills without the manpower and materiel costs associated with live-fire exercises. While simulation-based systems have long been leveraged by the Army as part of the continuum of Soldier training, their effectiveness is not often reliably assessed. A recent report by the Government Accountability Office (GAO 2013) found that neither the Army nor the Marine Corps sufficiently assess the effectiveness of its simulation training systems. To ensure the Army receives the best value for its investment, training effectiveness assessment (TEA) should be a priority for each training system the Army fields.

The Army and the Department of Defense (DOD) does conduct TEA to an extent, and there is evidence to suggest that simulation-based training is an effective means of providing Soldiers the learning experiences they need. However, the empirical research in this area is often mixed or inconclusive (Carretta and Dunlap 1998; Hays 2005). One reason for the ambiguity in these data are a lack of reliable and relevant performance metrics. Often, TEA is conducted through the administration of surveys to Soldiers, their instructors, or other observers. For example, a recent evaluation of the effectiveness of Virtual Battle Space 2 (VBS2) was conducted using self-report questionnaires about Warfighters' opinions of their own performance in the training scenarios (Ratawani et al. 2010). While promising, the validity of findings in studies like these is limited due to the inherently subjective nature of self-reporting.

A methodology is needed to develop objective metrics of Soldier performance within simulator systems. These metrics could be generated using the data simulators used to drive the training curriculum they provide. Currently, these metrics are not calculated, largely because training systems developers are not required to do so. In addition, there is no guidance for these developers with regard to how to identify the appropriate metrics for use in these systems. Research is needed to develop these guidelines and best practices to inform the overall training community going forward.

In addition to TEA, these metrics could inform a variety of audiences within the Army. For example, instructors and training developers in Army schoolhouses could use the data to track student performance, evaluate their instructors, and

compare courses over time. The metrics could be used to inform student models in adaptive training systems. Personnel involved in simulator maintenance could leverage the data to ensure their equipment is running consistently. Acquisitions personnel could use it to calculate return on investment, throughput, and other metrics of program success. Finally, research personnel could benefit from having the data available to them for experimentation.

Recent efforts from the US Army Research Laboratory (ARL) on Interoperable Performance Assessment (IPA) for individuals and teams have made advances on defining and persisting human performance data. The effort has concentrated on leveraging the work on the Experience application programming interface (xAPI) of the Advanced Distributed Learning (ADL) Initiative in an effort to produce data that have intersystem data value. The intent of the effort is to enable the creation of data across the training continuum that provides the following:

- Enables a historical view of proficiency;
- Demonstrates a live view of performance;
- Enables macro and micro adaptation; and
- Collects data for trends analysis for efficacy and effectiveness studies.

The effort has produced best practices and tools to encode performance data. The Soldier Performance Planner (SP²) and Pipeline are government off-the-shelf software tools that enable data to be captured and interacted with in interoperable ways. These tools were developed under a previous phase of IPA research that can be used to speed the time to interoperable data collection in other domains.

While capture and visualization capabilities exist, other challenges remain. Security, storage, and transmission of data, as well as general architecture, need to be explored to define long-term approaches. Additionally, challenges and solutions for identity management and protection of Personally Identifiable Information (PII) must be approached. Lastly, the means by which a Soldier's identity is determined at the network edge in ranges and simulator environments must be defined to further this work.

1.1.1 Goal

The Support for Training Effectiveness Assessment and Data Interoperability (STEADI) effort conducted the foundational research needed to provide guidance for how to develop effective measures and metrics for TEA, how to represent those measures and metrics in xAPI format, and how to design and develop an architecture to support their management; the research also provided guidance on

an interface so that instructors can easily access the data in meaningful ways. Additionally, the architecture would support persistence of the data for longer range studies.

1.1.2 Importance

Much of the future is focused on systems that provide an interoperable assessment capability, not only for individuals and teams, but for joint exercises, as well. Capturing individual and team performance data from multiple systems and connecting them in real time still poses a unique set of challenges that need to be explored further. The focus remains on having a highly agile system and approach that are capable of tracking experiences to improve the use of Live, Virtual, and Constructive (LVC) training for mission rehearsal. Such approaches are critical for future combat capabilities and threats in emerging environments.

2. Research Summary

2.1 Interoperable Performance Assessment (IPA)

The purpose of the previous research was to capture and leverage contextual performance measures from the xAPI to tailor learning to the individual learner's experience and competence level. Through the work performed, the concept of IPA was created and defined as "a method of uniformly defining and describing experience and context to assess learning and performance over time; to adapt training across a variety of environments, systems, and modalities, whereby performance is observed, assessed, evaluated, or asserted by systems or observers." This definition not only combines the methods in which interoperable tracking occurs, but also a) where the trainee, event, or training content is adapted and b) how the data are collected. While the definition remains broad, the intent is to move toward a common understanding of what is meant by interoperable tracking of performance data and the goals of assessing performance over time.

This capability of capturing and sharing performance data could provide the ability for systems to adapt and personalize learning experiences at both the micro and macro levels for future training events. Solutions in this domain provide an opportunity for cost and time savings, as performance data are shared between simulation, computer-based training systems, and other systems.

2.2 Human Performance Markup Language (HPML)

Current efforts that capture individual and collective performance include the HPML, an eXtensible Markup Language (XML) activity structure. This schema

was designed to capture and assess performance across distributed simulations by a language that identifies critical fields and stores them within an XML activity structure. The major goal in designing HPML was to focus on bridging the gap between the software implementation of measurements and assessments, and trainer design of the assessments. Overall, HPML was designed to allow the expression of important concepts from the training industry so that others, such as training professionals, instructors, operations, and researchers, can use, aggregate, and understand the data easily (Stacy et al. 2006). The current effort used HPML as a basis for describing current performance data that are being collected by various Army simulators, as well as to understand what type of system-based and observer-based data is being tracked across environments.

By leveraging analytics and metrics of performance-based activity data about individuals from a variety of sources, organizations can begin to provide the right support to unlock potential efficiencies. While adaptive and tailored opportunities for learning represent a path ahead to larger efficiency, there are currently shortcomings in the data availability across systems for interoperable tracking and assessment.

2.3 Experience API (xAPI)

Through the Training and Learning Architecture (TLA) effort, the ADL initiative is focusing on new approaches of standards for tracking learning experiences in a uniform way using the xAPI (Advanced Distributed Learning 2014). xAPI is a specification that allows an interoperable means to track experiences across Learning Management Systems (LMS), simulators, virtual worlds, web content, mobile devices, games, and observer-based measures. The project reached 1.0 specification in April 2013. Emerging technology known as a Learning Record Store (LRS) is being developed to store and track xAPI learning experiences. The LRS and GIFT are community-developed, open-source technology projects that allow the source code to be accessible by anyone for use or further development.

2.4 Soldier Performance Planner (SP²)

SP² is a tool for collecting and contextualizing Soldier performance data through web services and import tools. SP² is guided by IPA concepts and displays performance data of individuals and groups, which will lead to continuous training adaptation at macro and micro levels. SP² allows multiple data streams to be integrated into a single trainee profile. SP² also allows data for groups to be displayed. SP² has features that allow manipulation of data that includes import/export of data.

2.5 Pipeline

The “Pipeline” component for Microsoft.NET enables simulators to generate and consume training data to tailor the experience of learners in subsequent simulations, and, therefore, decrease the overall time required to complete training. Pipeline provides simulator vendors with a simplified software interface to track performance data during simulations using the xAPI specification, which are conformant to the “Encoding Best Practices Guide” as defined in the SP² project.

The Pipeline component reduces the complexity for software teams implementing the xAPI specification into their simulation products by abstracting many of the implementation details, such as transport and security. The component also enforces best practices by using a shared, common vocabulary so that performance data created during simulations can be used to tailor the experience in subsequent simulations for that same learner.

2.6 Current Research

STEADI encompasses one of several ongoing efforts focused on furthering investigation around IPA concepts using the xAPI. This effort conducted the foundational research needed to provide guidance for how to develop effective measures and metrics for TEA, how to represent those measures in the xAPI format, and how to design and develop an architecture to support their management; the research also provided guidance on an interface so that instructors can easily access the data in meaningful ways. Additionally, the architecture would support persistence of the data for longer range studies.

To complete this research, we leveraged a number of live and virtual training systems already in use by the Army, Marine Corps, and other agencies in the DOD. Specifically, data from the Modular Advanced Technologies Marksmanship Proficiency (MAT-MP) trainer and the location of misses and hits (LOMAH) sensors were used as test cases for this research.

During the research, we developed a long-term architecture that will guide future efforts of the services to broadly incorporate IPA concepts across the continuum of training.

Throughout the course of this effort, we engaged the end-user community through workshops, conference presentations, and publications to help develop a community of practice for ARL’s vision for data interoperability standards.

3. Approach

3.1 Project Goals

STEADI focused on evaluating the effectiveness of encoding data through xAPI, and sharing data across trainers and other systems within a learning ecosystem. Specifically, the object was to develop cross-platform training effectiveness measures and metrics for the marksmanship training domain using the xAPI.

The primary goals of STEADI were to conduct the foundational research to guide how to develop effective measures and metrics for TEA, and represent those measures and metrics in xAPI format. Additionally, this effort aimed to design and develop an architecture to support the management of TEA measures and metrics, using the xAPI, with an interface enabling easy and meaningful data access for instructors. Moreover, this architecture was aimed to support the persistence of data in alignment with IPA concepts and objectives.

Technical objectives of the effort included the following:

- User needs analysis of marksmanship training
- Development of measures for marksmanship
- Instantiation of measures in xAPI format
- Design of an architecture to manage the data and automatic measure development
- Conduct validation of measures and architecture in TEA
- Analyze and report on data collected for the evaluation effort

3.2 Project Tasks

3.2.1 Task 1: User Needs Analysis

Task 1 focused on identifying the needs of the marksmanship training community, including researchers, training developers, instructors, and resource managers, as they relate to improved access to marksmanship data. This was accomplished through a user needs analysis, which included a literature review and data collection with representatives of our end-user communities.

3.2.1.1 Literature Review

To familiarize ourselves with the issues involved in Army basic rifle marksmanship and potential measures of performance, we examined journal articles, technical reports, and conference proceedings from databases including ESBCO (Academic Search Complete, Military and Government Collection, PsycArticles), Defense Technical Information Center (DTIC), and Google Scholar. In addition, we reviewed proceedings from conferences including the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC) and Human Factors and Engineering Symposium (HFES), among others. This literature review focused on 3 areas: 1) previous assessments of training effectiveness of marksmanship training technology; 2) the knowledge, skills, and abilities (KSAs) relevant to marksmanship performance; and 3) the ways shooter skills can be directly assessed.

From this literature review, we were able to develop interview and focus group protocols for implementation. Importantly, we identified a model of skill development to serve as the basis for our measure development going forward. The research literature suggests marksmanship performance is dependent upon cognitive, psychomotor, and affective components. Our review summarizes these components and the relative contributions they have on marksmanship performance at different points throughout the periods of instruction. Based on this model, we were able to identify prototype measures of marksmanship performance.

3.2.1.2 Interviews and Focus Groups

Semistructured interview and focus group protocols were developed to guide the discussions, and were provided to ARL for review. The questions centered around issues with designing, delivering, and managing marksmanship training, although the specific content varied based on the position of the interviewee. Researcher interviews focused on their personal and team research goals, issues involving conducting marksmanship research, and the potential research questions that could be answered with improvements in data collection and management. For individuals directly involved in marksmanship training, including instructors, the questions dealt with how performance was diagnosed, the typical errors shooters make, and remediation strategies. Interview protocols with resource management personnel focused on the processes underlying supplying trainers with the necessary ammunition, range, and simulation time. Training developer interviews included questions about how programs of instruction (POIs) are generated, how marksmanship training is evaluated, and how emerging training strategies are being developed and deployed.

To implement these protocols, targeted user communities were identified by ARL personnel, and interviews and focus groups were coordinated through Maneuver

Center of Excellence (MCoE) leadership. The data collection was conducted from September 28, 2015 to October 2, 2015 at Fort Benning, Georgia. Interviews and focus groups were conducted with Research psychologists from the US Army Research Institute for the Behavioral and Social Sciences (ARI), Engagement Skills Trainer (EST) proponent instructors, 194th Armor Brigade (AR BDE) instructors, Marksmanship Master Trainer Course (MMTC) trainers, training and task developers from the MCoE Department of Training Development (DOTD), Simulation center managers, range control operations personnel, ammunition resource managers, and drill sergeants.

In addition to the execution of the interview and focus group protocols, subject matter expert (SME) feedback was collected on the MAT-MP system, under development by the Naval Air Warfare Center–Training Systems Division (NAWC–TSD). To collect these data, an instructor interface demonstration was conducted and the system was described. The findings from this process were provided to ARL in a separate report to inform the best ways this system could be used in the training process.

3.2.1.3 Data Analysis and Reporting

The resulting data from the user needs analysis were collected, combined, and analyzed in order to determine the following:

- What are the primary issues involved in conducting and managing marksmanship training for each user group?
- What are the cognitive, psychomotor, and affective components of learning to shoot, and how are these assessed?
- How are issues in each of these components diagnosed and remediated?
- How could the marksmanship training process be improved through data analytics?
- How are resources managed, and how could this process be improved through improved access to data?
- How is training effectiveness currently assessed?
- How could research methods improve as a result of increased data accessibility?

When taken together with the findings from the literature review, we found considerable support for our model of marksmanship skill acquisition.

Additionally, new constructs of skills related to marksmanship were identified for inclusion in our measure battery.

3.2.2 Task 2: Measure Development

Task 2 focused on developing the relevant measures of marksmanship performance for inclusion in the STEADI architecture. Findings from the literature review and user needs analysis completed in Task 1 were leveraged to develop preliminary measures. While the focus measures revolved around TEA, candidate measures also included those required by other audiences, such as resource management or research.

3.2.2.1 Identification of Measures

Based on our literature review and SME interviews and focus groups, relevant measures of marksmanship performance were determined. These measures were based on the theoretical support of their predictive effectiveness, predictive validity, and ease of implementation. While pilot data on these measures has yet to be collected, our findings suggest the inclusion of the measures described in Table 1. Measures of these constructs were collected and combined into a prototype assessment battery. The test battery is designed to be comprehensive as an initial research tool; further research validating our model will determine the extent to which these tests should be included, modified, or removed.

In addition to these measures, our team identified measures of marksmanship performance that could serve as criteria in validating this model going forward. These include qualification scores, range system data, sensor data, shot analysis, and guided instructor assessments. To support data collected directly from instructors, a paper-based manikin assessment was developed. This tool helped instructors to quickly identify specific body position problems, which acted as a means of collecting data while also providing an easy-to-use tool for remediation. The content of the tool was based on previous, checklist-based tools (James and Dyer 2011). While these checklists have proven useful in the field, our goal was to develop a measure that could provide easy to code data for inclusion in our model. This measure was well-received during our user needs analysis. As a result, this tool was reviewed with ARL and integrated into the STEADI architecture in Task 4.

Table 1 Rifle marksmanship assessment constructs

Component	Construct
Cognitive	General cognitive ability Marksmanship domain knowledge Openness to experience Hunting experience Videogame experience
Psychomotor/physical	Visual acuity Handedness Eye dominance Height Physical fitness Sports experience Musical ability
Affective	Perceived stress Resiliency/hardiness Grit Self-efficacy Initiative

3.2.2.2 Evaluation Study Design

We designed a research strategy to demonstrate several data-related use cases over the course of a 2-week training cycle. By answering a broad number of research questions, we aim to demonstrate the value of implementing performance measurement across the marksmanship training process. A full description of the study design was provided in the user needs analysis report, but the research questions we aim to address are identified as follows:

- How is marksmanship knowledge developed and retained during the training process?
- Does marksmanship knowledge and/or measures of cognitive, psychomotor, and affective constructs predict the ability to group and zero?
- Is the sensor data collected in the MAT-MP system consistent with that collected in the EST II?
- Is the MAT-MP consistent with instructor assessments of performance?
- How does the relative contribution of cognitive, psychomotor, and affective components change over the course of training?
- To what extent do knowledge tests, our assessment battery, and the MAT-MP predict performance on a known distance range?
- To what extent do all these measures predict final qualification scores?

The schedule for data collection is presented in Table 2. Importantly, we propose to conduct this research with the same Soldiers throughout the training cycle.

Table 2 Proposed data collection schedule

BRM period	Measures
BRM 1	Domain knowledge Cognitive/psychomotor/affective measures
BRM 2/3	Instructor assessment EST II MAT-MP
BRM 6	Domain knowledge Cognitive/psychomotor/affective measures LOMAH/shot accuracy MAT-MP Instructor assessment
BRM 13	Qualification score

3.3.3 Task 3: Instantiation of xAPI Metrics

Task 3 focused on integrating data capture to support the measures identified in Task 2. The measures were integrated into an xAPI registry for use in the data encoding library component, Pipeline, and marksmanship training systems. The outcome of this task was the capability to enable extraction of xAPI statements from specified systems via connection to the xAPI registry.

3.3.3.1 Method

Originally, this task was focused on updating the data encoding library, Pipeline, with the measures identified in Task 2. However, this approach was adjusted due to lack of access to target training systems (e.g., EST). Without access, the data formats of each system were unknown. There was a great risk for updating the current version of Pipeline and finding the target systems were incompatible. Likewise, an update of Pipeline would limit the flexibility for incorporation of varied training systems within the STEADI architecture. In order to maximize plasticity and minimize reengineering requirements, an online xAPI registry was created to include the identified marksmanship measures.

The STEADI xAPI registry approach provides the necessary legwork to update Pipeline in the future, if necessary. It also acts as a standalone reference for training simulator vendors to easily incorporate STEADI-specific, xAPI measures. In addition, the online registry provides a central means to quickly update the core measures and metrics associated with STEADI during future research efforts.

Prior to developing the online registry, the measures identified in Task 2 were documented in a spreadsheet. All applicable information was included, such as name, description, source, or xAPI property type. This matrix facilitated rapid development of the online registry. Table 3 provides a sample matrix with marksmanship measures, formatted the same as the one created for Task 3.

Table 3 Sample marksmanship measures for Pipeline update

Measure	Description	Data type	Data range	Source	xAPI property
Team sport experience	Team sport participation frequency	Integer	1–5; not at all to very frequently	Demographics Questionnaire	Context
Eye side dominant	Eye dominance: left or right	String	(a) left eye, or (b) right eye	Demographics Questionnaire	Context
BRMT	Basic rifle marksmanship test score	Integer	0–10	Basic rifle marksmanship testing score	Context

The STEADI xAPI registry was developed in a webpage to increase accessibility. An initial website design identified the visual (i.e., layout) and functional requirements. The online registry was created using Hypertext Markup Language (HTML), Cascading Style Sheets (CSS), and JavaScript. The layout of the page was designated by CSS, while the actual information within the page, including the primary table with the marksmanship measures, was made using HTML. JavaScript was leveraged to enable particular behaviors for the site. Specifically, adding a script allowed the ability to click the “Example” button for a specific measure and seamlessly jump to the appropriate location to view the sample xAPI statement for that measure. Once clicked, the example was designed to appear at the top of the webpage. A buffer was also added to prevent the example from hiding behind the static banner located at the top of the page.

3.3.3.2 Results

An online xAPI code registry incorporating the identified marksmanship measures was designed and created based on the measure matrix. Figures 1 and 2 capture a sample of the level of detail included in the registry. The information provided will facilitate the process of implementing xAPI in other systems.

Team Sport Experience	Team sport participation frequency.	http://www.problemsolutions.net/xapi/contextextensions/team-sport-experience	Context	Example
Individual Sport Experience	Individual sport participation frequency.	http://www.problemsolutions.net/xapi/contextextensions/individual-sport-experience	Context	Example
Large Animal Hunting	Large game hunting frequency.	http://www.problemsolutions.net/xapi/contextextensions/large-animal-hunting	Context	Example
Small Animal Hunting	Small game hunting frequency.	http://www.problemsolutions.net/xapi/contextextensions/small-animal-hunting	Context	Example
Fishing	Fishing frequency.	http://www.problemsolutions.net/xapi/contextextensions/fishing	Context	Example
Outdoor Activities	Time spent outside frequency.	http://www.problemsolutions.net/xapi/contextextensions/outdoor-activities	Context	Example
Media Consumption	How much media is used on a regular basis.	http://www.problemsolutions.net/xapi/contextextensions/media-consumption	Context	Example

Fig. 1 STEADI xAPI registry: main table

Team Sport Experience

Description: Team sport participation frequency.

```
"context": {
  "extensions": {
    "http://www.problemsolutions.net/xapi/contextextensions/team-sport-experience": 3
  }
}
```

Individual Sport Experience

Description: Individual sport participation frequency.

```
"context": {
  "extensions": {
    "http://www.problemsolutions.net/xapi/contextextensions/team-sport-experience": 3
  }
}
```

Fig. 2 STEADI xAPI registry: example xAPI statements

3.3.4 Task 4: Architecture Design

In Task 4, we developed an architecture for IPA data capture and usage. The project team mapped key use cases for data capture, identity management, storage, secure access, and data analysis. The use cases were employed to create system architecture designs. The design considered current training and records management systems, and planned systems to move to the end state. The literature review and user needs analysis completed in Task 1 were also leveraged to inform the design of the architecture. Task 4 provided output that will assist to describe and procure technical solutions to fulfill the architecture use cases across military systems.

3.3.4.1 Method

Use Case Analysis

The Unified Modeling Language (UML) was leveraged to design, describe, and visualize aspects of the STEADI architecture. The UML provides a means to describe requirements and design intent, and a guide for development, as well as

the facilitation of reverse engineering and documentation (International Organization for Standardization 2005). The UML use case model and diagram was employed to capture the requirements of the STEADI architecture. Use cases provided a means of communicating what the system(s) are intended to do. Moreover, the delineation of key use case requirements served as building blocks for the final architecture design. Figure 3 shows a draft version of the STEADI use case diagram.

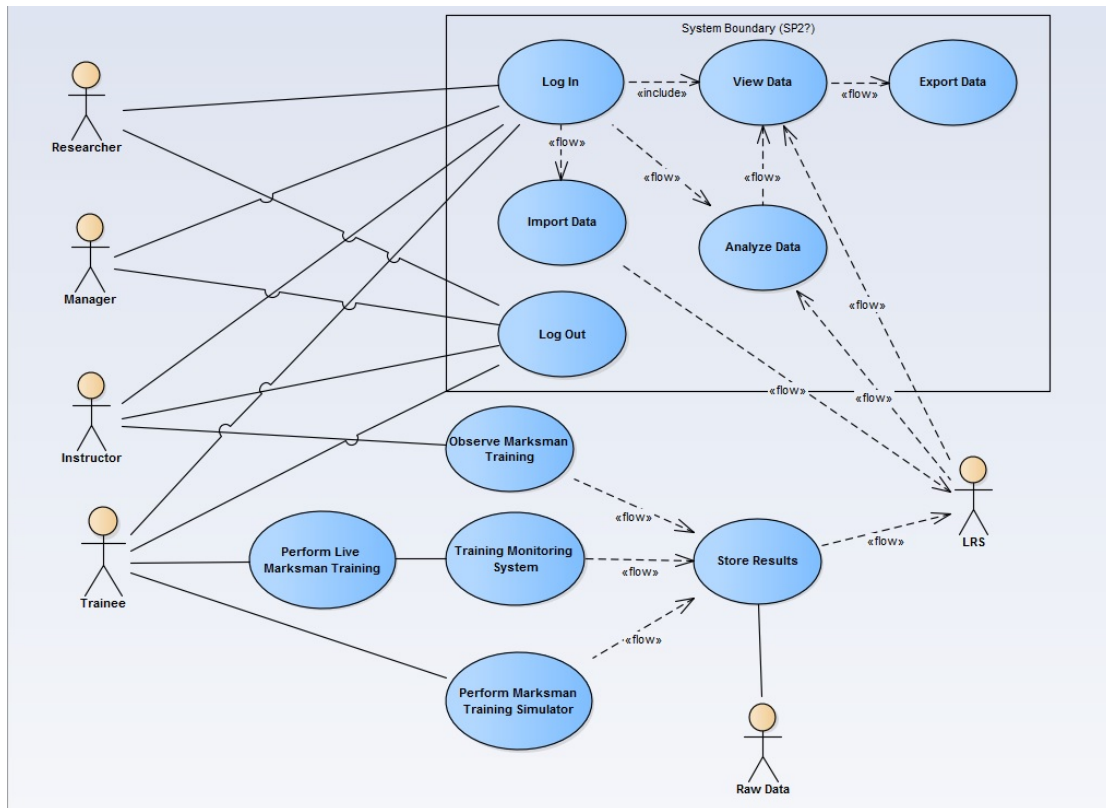


Fig. 3 STEADI use case diagram working draft

Key use cases were mapped for data capture, identity management, storage, secure access, and data analysis. Each use case was analyzed to determine specific requirements. Actors, entities external to the architecture (e.g., instructors), were also identified to inform and organize the use case requirements. Leveraging this use case analysis, potential architecture solutions, focused on open-source technologies, were identified.

A use case diagram was designed to display the interactions between the system and actors. This diagram provided a visualization for all the mapped use case requirements. In addition, the primary use cases and actors were reviewed and detailed in a use case document.

Architecture Design

Enterprise Architect, a UML architecture tool developed by Sparx Systems Pty Ltd., was used to develop both the use case and architecture diagrams. The STEADI architecture was developed using the use case analysis and diagrams. The requirements visualized in the use case diagrams provided the basis to develop the overarching system diagram.

Figure 4 displays a visualization of the method and Enterprise Architect tool used to design and develop the STEADI architecture. The primary use cases can be seen on the right, whereas the architecture design is on the left. Note that the displayed diagram is not the final architecture design.

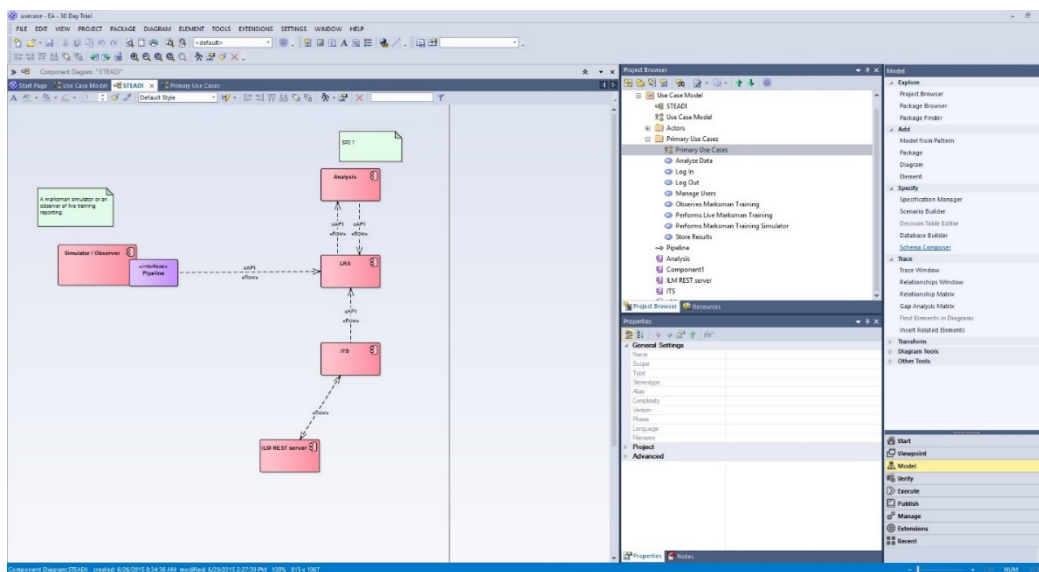


Fig. 4 STEADI Architecture Design Methodology and Tool

3.4.4.2 Results

Use Case Analysis

The results of the use case analysis were visualized in an overarching use case diagram. The final use case diagram is shown in Fig. 5. The actors—systems or users external to the central system(s)—are shown as stick figures. The main use case requirements, or the functional requirements, including the data flows, are visualized in the use case diagram.

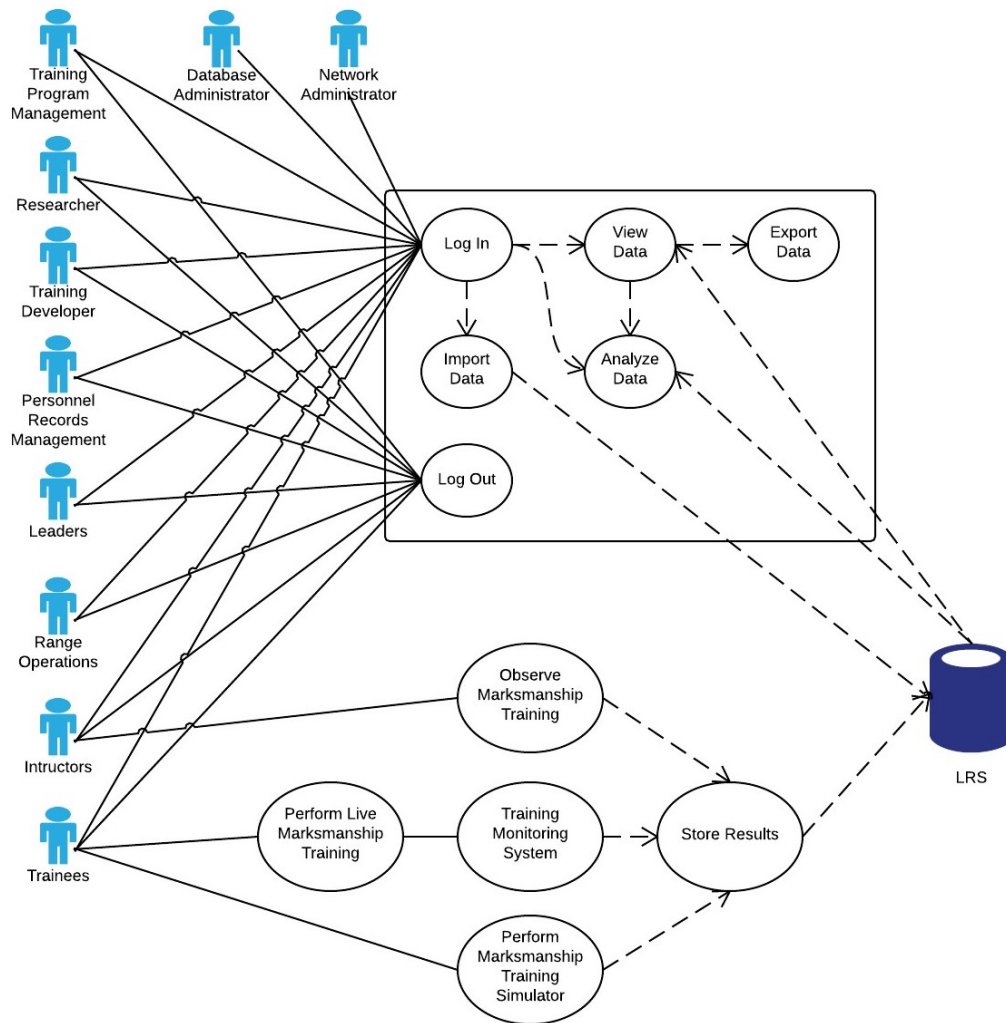


Fig. 5 STEADI use cases diagram

In addition to the STEADI Use Cases Diagram, a use case description was developed. All the use cases, actors, requirements, scenarios, constraints, and other attributes are included in the use case document.

Architecture Design

The final STEADI architecture design was visualized in an architecture diagram (shown in Fig. 6). The architecture design meets all use case requirements, and depicts the system boundaries, data flows, scenarios, and other attributes. The diagram provides a system-level overview of the STEADI architecture, which will provide guidance for developing a reference architecture in following efforts.

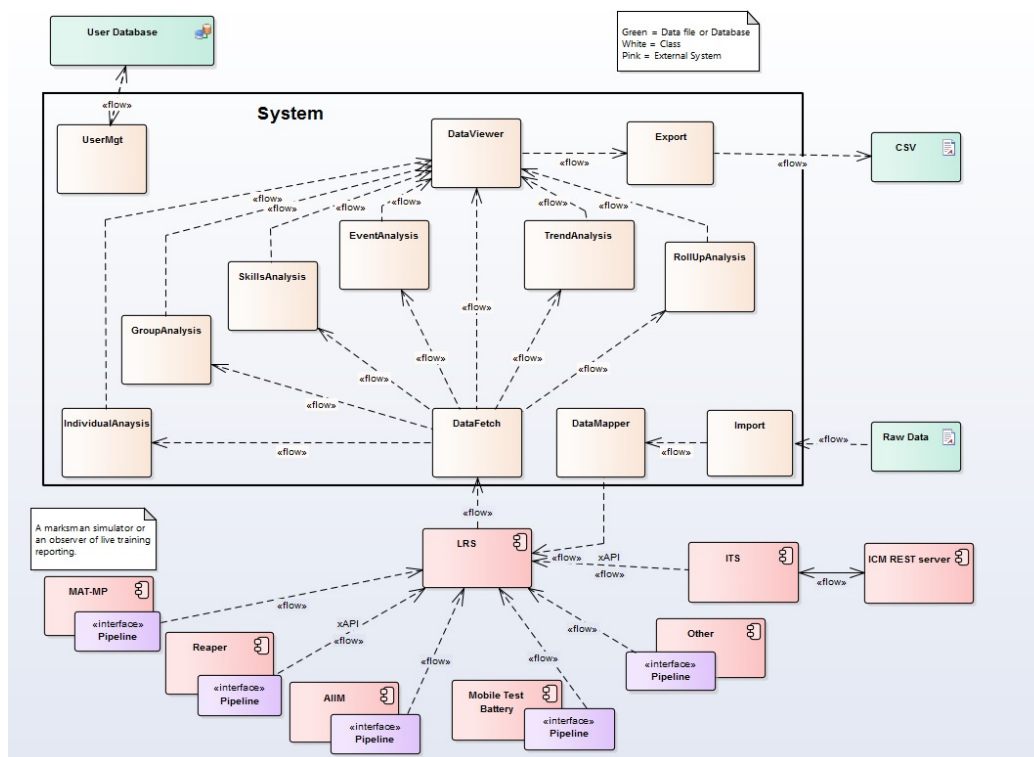


Fig. 6 STEADI architecture diagram

4. Results and Discussion

4.1 Overall Findings

A comprehensive literature review and user needs analysis was conducted to inform a variety of research questions. Interestingly, we found that Army marksmanship training was in a transitory state. Both the training systems and doctrine were being updated during the course of this effort. EST, for instance, was mostly inaccessible during our research. The system was being replaced with an upgraded model as a result of a contract award to a new EST vendor. As such, we were unable to access EST for testing or data collection. Conversely, the MAT-MP system resided in research and was not fielded at the time of our study; however, we were able to collect feedback from SMEs during our user needs analysis. We found a high level of interest among the user communities and developed a separate report summarizing this review.

As a result of the literature review and user needs analysis, we developed preliminary marksmanship measures for TEA. While these findings require additional validation, they point to the potential for a wide range of measures that may impact marksmanship performance. Beyond typical psychomotor measures

for marksmanship, supplemental cognitive and affective constructs (e.g., hunting experience or grit) may impact training and performance outcomes.

An xAPI registry was also developed to instantiate the identified measures in xAPI. This online registry enabled us to mitigate data access issues for marksmanship training systems. In comparison to updating Pipeline, the registry also offered more flexibility, eased future update requirements, and minimized potential reengineering risks.

Finally, an architecture was designed to support the persistent management of marksmanship measures and metrics for TEA. We identified multiple user communities during an initial use case analysis. In addition to the primary trainee/trainer user groups, our analyses revealed additional communities, such as training managers, personnel management, and range operators. These secondary user communities were not originally hypothesized to gain value from accessing the STEADI architecture. However, the information supported through the developed architecture may prove to have impact beyond our initial, core user communities.

Several open source tools and technologies were identified as having potential to support development of a STEADI reference architecture. Primarily, SP² was recognized as a prime candidate for the STEADI user interface. It currently provides a number of the functional requirements identified during the use case analysis. While additional development would be required, SP² offers a good starting point to accelerate development efforts and lower costs to ARL. Acting as the central glue between all the systems, the xAPI specification was selected early on as the prime candidate to enable data interoperability and persistence across different systems. As such, the STEADI xAPI registry and Pipeline serve to enable systems within the STEADI architecture to produce uniform xAPI statements, which conform to the identified marksmanship measures.

4.2 Future Recommendations

The findings of our research so far are promising; however, additional validation is required to determine the validity of the initially identified marksmanship measures. We have designed a research strategy to validate the measures developed under this effort, as well as to answer a broad range of research questions regarding the time course of the development of cognitive, affective, and psychomotor aspects of marksmanship training. Once the measures are validated, a next step would be to validate the metrics of these measures to distinguish expert from novice performance.

Metrics, in turn, will require development and validation. Leveraging a number of live and virtual environments for study will enable the process for evaluating the validity of the metrics, as well as measures, being incorporated into the STEADI architecture. Importantly, collecting performance data across multiple periods of marksmanship instruction, using multiple criteria, will support an understanding of how performance changes throughout the training process.

During our interviews, several instructors supported the concept of developing a mobile application to support performance evaluation. This application would provide instructors the opportunity to photograph or video a trainee as they take a shot and compare the position and movements of the shooter with those of an expert. Such an application would enable researchers to gather instructor evaluations of trainees' performance, while providing instructors a valuable training aid.

In the context of this mobile platform, we recommend developing tablet-based versions of our measures for more efficient data collection and management. In addition, measures of psychophysical skills, such as reaction time and visuospatial processing, could be included through third-party software.

During our User Needs Analysis, marksmanship instructors provided a variety of opinions regarding the usefulness of the MAT-MP as an effective training tool for basic rifle marksmanship. Several instructors suggested the system, when developed, be evaluated by them for incorporation into the marksmanship curriculum. In order to determine the appropriate use for this system, an evaluation should be conducted.

Currently, the MAT-MP provides data regarding trigger pressure, cant, buttstock pressure, and a video feed of the shooter's sight picture. The extent to which these measures provide actionable information to instructors has yet to be determined. Future research should develop reliable and valid metrics of performance using this system, as well as diagnostic criteria and a remediation framework.

Further investigation into the use of an xAPI registry for marksmanship should be completed. Cross-platform access to identified marksmanship measures and metrics must take in to account the potential for differing data formats. The current registry requires additional update upon measure validation and metric identification, as well.

Several tools and technologies that were incorporated into the STEADI architecture design are recommended for further review and inclusion in future research:

- Mobile Test Battery: A combination of standard and custom tests to support comprehensive data collection and evaluation of cognitive, psychomotor,

and affective measurements for marksmanship. A mobile capability increases accessibility in diverse training locations and offers the ability to incorporate result tracking using the xAPI specification.

- Assessment Instructor Interface for Marksmanship (AIIM): An interactive, mobile marksmanship support tool for live responsive instructor assistance for assessing shooter problems. AIIM is the recommended next step to the paper-based, manikin assessment tool created during the user needs analysis. Creating a mobile application affords significant tool functionality and data accessibility via xAPI. By leveraging tablets, AIIM will allow instructors to easily take “live” pictures of shooters for simple, personalized problem remediation support. In addition, instantiation of these data in xAPI would enable incorporation into the STEADI architecture.
- SP²: The primary user interface to allow end-users to access and interpret marksmanship data within the STEADI architecture. Since SP² already incorporates some of the desired functionality and is a government off-the-shelf tool, it provides the ideal platform to create the STEADI user interface. Additional review and update are recommended for incorporation of identified STEADI architecture requirements.
- xAPI Registry/Pipeline: The central, cohesive component to the architecture that supports data interoperability and persistent management of the marksmanship measures and metrics for TEA across different systems. Additional investigation on the use of the xAPI specification within the STEADI architecture via Pipeline and the xAPI Registry is recommended.

5. Conclusion

The research conducted under this effort provides an important step toward demonstrating the benefits of improving performance measurement and data accessibility in Army training. Using a marksmanship use case, our research lays the groundwork for developing and implementing measures to improve the effectiveness and efficiency of the training process.

Developing the STEADI architecture has led to the identification of challenges associated with the use of xAPI to store trainee performance measures. For example, something that is both the biggest strength and the biggest weakness of xAPI is that it does not prescribe how particular performance measures are to be described. It is a strength because it creates flexibility for systems that would use it to save user activities. Conversely, it is a weakness because different systems can

end up encoding the same behavior in several different ways, making it challenging to compare performance across systems. ADL has relied on communities of practice to solve this problem. Communities of practice are expected to develop their own standards for encoding specific behaviors to include specific naming conventions and the like. The creation of an xAPI registry, as described previously, provides a means to catalogue those conventions in a tool that facilitates the generation of standard xAPI statements.

Although the xAPI registry establishes common ways of describing trainee behaviors using xAPI statements, it does not solve the problem of 2 different systems measuring the same behavior in different ways. For example, the Engagement Skills Trainer, a simulator, measures trigger pull by looking at movement of the trigger. On the other hand, a device used to record trigger pull during live-fire training measures trigger pressure. While both of these systems measure trigger pull, the sensors used would not be expected to produce equivalent output. One solution to this problem would be to have individuals of different ability levels produce comparable trigger pulls on both systems and then attempt to develop a transformation function to produce comparable output. Another would be to treat these as different but related measures that each feed into diagnostic and/or predictive models, each one being validated independently. These solutions will have to be evaluated in future research.

This leads to the final challenge of using xAPI to record data about the learner: raw performance measures are frequently not informative in and of themselves. In addition to the need for supplemental contextual data, as discussed above, models are often needed to provide metrics that an instructor can interpret. Using the trigger pull example just described, a graph of pressure change over time or even trigger movement over time would not make sense to even an experienced instructor. Models are needed help the instructor interpret whether the data represents a proper trigger pull for a student at that point in the training program. As we proceed with our research, we will be investigating the development of these models to facilitate interpretation of data and to determine the best way to present that data to trainers and other user groups.

6. References

- Advanced Distributed Learning. Training and Learning Architecture (TLA). Retrieved from Advanced Distributed Learning [accessed May 2015]. <http://www.adlnet.gov/capabilities/tla>, 2014.
- Carretta TR, Dunlap RD. Transfer of training effectiveness in flight simulation: 1986 to 1997. Mesa (AZ): US Air Force Research Library, Human Effectiveness Directorate; 1998.
- Durlach PJ, Ray JM. Designing adaptive instructional environments: insights from empirical evidence. Arlington (VA): Army Research Institute Report; 2011. Technical Report 1297.
- GAO. Army and Marine Corps training: better performance and cost data needed to more fully assess simulation-based efforts. Washington (DC): US Government Accountability Office, 2013.
- Hays RT. The effectiveness of instructional games: A literature review and discussion. Orlando (FL): Naval Air Warfare Center Training Systems Division; 2005. Report No.: NAWCTSD-TR-2005-004.
- International Organization for Standardization. Unified modeling language specification: version 1.4.2. Object Management Group [accessed November 2015]. www.omg.org/spec/UML/, 2005 January.
- Landsberg CR, Astwood RS, Van Buskirk WL, Townsend LN, Steinhauser NB, Mercado AD. Review of adaptive training system techniques. *Military Psychology*. 2012;96–113.
- Ratwani KL, Orvis K, Knerr BW. Game-based training effectiveness evaluation in an operational setting. Arlington (VA): US Army Research Institute for the Behavioral and Social Sciences; 2010. Research Report (655803).
- Stacy EW, Ayers J, Freeman J, Haimson C. Representing Human Performance with Human Performance Measurement Language. Proceedings of the Fall 2006 Simulation Interoperability Workshop, Orlando (FL).
- US Department of the Army. The US Army learning concept for 2015. Washington (DC): Department of the Army; 2011. TRADOC Pam 525-8-2.

List of Symbols, Abbreviations, and Acronyms

ADL	Advanced Distributed Learning
AIIM	Assessment Instructor Interface for Marksmanship
API	application programming interface
AR BDE	Armor Brigade
ARI	Army Research Institute for the Behavioral and Social Sciences
ARL	Army Research Laboratory
CSS	Cascading Style Sheets
DOD	Department of Defense
DOTD	Department of Training Development
DTIC	Defense Technical Information Center
EST	Engagement Skills Trainer
HFES	Human Factors and Engineering Symposium
HPML	Human Performance Markup Language
HTML	Hypertext Markup Language
I/ITSEC	Interservice/Industry Training, Simulation and Education Conference
ID	identification
IPA	Interoperable Performance Assessment
KSAs	knowledge, skills, and abilities
LMS	Learning Management System
LOMAH	Location of Misses and Hits
LRS	Learning Record Store
LVC	Live, Virtual, and Constructive
MAT-MP	Modular Advanced Technologies Marksmanship Proficiency
MCoE	Maneuver Center of Excellence
MMTC	Marksmanship Master Trainer Course

NAWC–TSD	Naval Air Warfare Center Training Systems Division
PII	Personally Identifiable Information
POI	Program of Instruction
SME	subject matter expert
SP ²	Soldier Performance Planner
STEADI	Support for Training Effectiveness Assessment with Data Interoperability
TEA	Training Effectiveness Assessment
TLA	Training and Learning Architecture
UML	Unified Modeling Language
US	United States
VBS2	Virtual Battle Space 2
xAPI	Experience API
XML	eXtensible Markup Language

1 DEFENSE TECHNICAL
(PDF) INFORMATION CTR
DTIC OCA

2 DIRECTOR
(PDF) US ARMY RESEARCH LAB
RDRL CIO L
IMAL HRA MAIL & RECORDS
MGMT

1 GOVT PRINTG OFC
(PDF) A MALHOTRA

2 PROBLEM SOLUTIONS INC
(PDF) MIKE HRUSKA
KRISTIN MAHLA

1 QUANTUMIMPROVEMENTS
(PDF) LLC
JENNIFER MURPHY

1 ARMY RSCH LAB – HRED
(PDF) RDRL HRB B
T DAVIS

7 ARMY RSCH LAB – HRED
(PDF) SFC PAUL RAY SMITH
CENTER
RDRL HRO COL H BUHL
RDRL HRA I MARTINEZ
RDRL HRR R SOTTILARE
RDRL HRA C RODRIGUEZ
RDRL HRA B J HART
RDRL HRA A C METEVIER
RDRL HRA D B PETTIT

1 USA ARMY G1
(PDF) DAPE HSI B KNAPP

1 USAF 711 HPW
(PDF) 711 HPW/RH K GEISS

1 USN ONR
(PDF) ONR CODE 34 P MASON

1 USN ONR
(PDF) ONR CODE 341 J TANGNEY

1 OSD OUSD ATL
(PDF) HPT&B B PETRO

1 USA NSRDEC
(PDF) RDNS D D TAMILIO

ABERDEEN PROVING GROUND

11 DIR USARL
(PDF) RDRL HR
P FRANASZCZUK
RDRL HRB
J LOCKETT
RDRL HRB A
M LAFIANDRA
RDRL HRB C
J GRYNOVICKI
RDRL HRB D
D HEADLEY
RDRL HRF
K OIE
RDRL HRF A
A DECOSTANZA
RDRL HRF B
A EVANS
RDRL HRF C
J GASTON
RDRL HRF D
A MARATHE
RDRL HRT T
G GOODWIN

INTENTIONALLY LEFT BLANK.