Developing Interoperable Data for Training Effectiveness Assessment in Army Marksmanship Training

Jennifer S. Murphy
Quantum Improvements Consulting
Orlando, FL
jennifer.murphy@quantumimprovements.net

Michael Hruska
Problem Solutions
Johnstown, PA
mike@problemsolutions.net

Gregory A. Goodwin; Charles R. Amburn
U.S. Army Research Laboratory – Human Research and Engineering Directorate
Orlando, FL
gregory.a.goodwin6.civ@mail.mil; charles.r.amburn.civ@mail.mil

ABSTRACT

As Army training budgets become more limited, a renewed focus has been placed on simulation-based training to maintain Soldier readiness. 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 their simulation training systems. One barrier to conducting this assessment is the inability to leverage objective metrics of Soldier performance within simulation-based and live marksmanship training to optimize training outcomes. This paper discusses the foundational research being conducted to develop an architecture that will allow metrics for marksmanship training collected using the Experience API (xAPI) format, to support training effectiveness analysis, adaptive training systems, and training resource management.

ABOUT THE AUTHORS

Jennifer Murphy is the Chief Executive Officer of Quantum Improvements Consulting, LLC. She has over 10 years of military selection and training research experience, with an emphasis on leveraging innovative technologies for improving training in a measurably effective way. Upon completion of her Ph.D. from the University of Georgia in 2004, Dr. Murphy took a position as a Research Psychologist at the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI). Her research focused on the development of technology-based selection and training measures for cognitive and perceptual skills. Dr. Murphy served as Director of Defense Solutions at Design Interactive, Inc., where she managed a portfolio of training and performance support efforts incorporating cutting edge technology into training solutions for defense clients. Her research has been featured in *The New York Times*, the Pentagon Channel, *Soldier* Magazine, and *Signal* Magazine.

Michael Hruska is a technologist with experiences spanning across standards, emerging technologies, learning, and science. He is a former researcher at the National Institute of Standards and Technology in Gaithersburg, Maryland. He is currently the President / Chief Executive Officer of Problem Solutions, and provides learning technology solutions to government, commercial, and nonprofit organizations. His team has been supporting efforts for the last 4 years at the Advanced Distributed Learning (ADL) Initiative on the future of a Training and Learning Architecture (TLA) and the Experience Application Programming Interface. He holds a Bachelor of Science from the University of Pittsburgh and is a member of the e-Learning Guild, American Society of Training and Development (ASTD) and the National Defense Industrial Association (NDIA).

Gregory Goodwin is a senior research scientist at the Army Research Laboratory-Human Research and Engineering Directorate, Simulation and Training Technology Center (STTC) in Orlando, Florida. His research focuses on methods and tools to maximize the effectiveness of training technologies. After completing his Ph.D. at the State University of New York at Binghamton in 1994, Dr. Goodwin spent three years in a post-doctoral fellowship at the

Columbia University College of Physicians and Surgeons followed by a year as a research associate at Duke University Medical Center before joining the faculty at Skidmore College. In 2005, Dr. Goodwin left academia and began working at the Army Research Institute (ARI) field unit at Fort Benning Georgia and six years later, he came to the ARI field unit in Orlando, FL where he has been examining ways to leverage technologies to reduce the cost and improve the effectiveness of training.

Charles Amburn is the Senior Instructional Systems Specialist for the United States Army Research Laboratory, Human Research and Engineering Directorate, Simulation Training and Technology Center (STTC) in Orlando, Florida. After obtaining both a Film degree and a Master's degree in Instructional Systems Design from the University of Central Florida, he began his Department of Defense civilian career in the Advanced Instructional Systems Branch at the Naval Air Warfare Center Training Systems Division (NAWCTSD). There he worked on special projects for the Navy and Marine Corps for 10 years before becoming the Lead Instructional Designer for the Army's Engagement Skills Trainer (EST) program at the Program Executive Office for Simulation, Training and Instrumentation (PEO-STRI), Orlando, Florida. Since 2011, Mr. Amburn has worked on various projects within the Soldier-Centered Army Learning Environment (SCALE) and Joint and Coalition Training Rehearsal and Exercise Research (JCTRER) research programs at STTC.

Developing Interoperable Data for Training Effectiveness Assessment in Army Marksmanship Training

Jennifer S. Murphy
Quantum Improvements Consulting
Orlando, FL
jennifer.murphy@quantumimprovements.net

Michael Hruska
Problem Solutions
Johnstown, PA
mike@problemsolutions.net

Gregory A. Goodwin; Charles R. Amburn
U.S. Army Research Laboratory – Human Research and Engineering Directorate
Orlando, FL
gregory.a.goodwin6.civ@mail.mil; charles.r.amburn.civ@mail.mil

INTRODUCTION

The U.S. Army is returning from over a decade at war to unprecedented training challenges. Going forward, the Army faces a hybrid threat of conventional and unconventional warfare. Maintaining combat overmatch in this ever-changing environment requires innovative and effective training solutions. At the same time, the Army faces smaller training and operational budgets. Key to the Army's success in developing and maintaining adaptable, ready Soldiers will be to maintain training quality leveraging the newest, most innovative technologies within these financial constraints.

One way of maximizing training efficiency is the use of simulation-based training. Simulation, if leveraged effectively, can save live trials, increase throughput, reduce instructor workload, and shorten training time. While simulation-based training has long been used by the Army as part of its training curriculum, the effectiveness of these systems 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 their simulation training systems. To ensure the Army receives the best value for their investment, Training Effectiveness Assessment (TEA) should be a priority for each training system the Army fields.

The measurement of training effectiveness should be a continuous process throughout the lifespan of a training simulator. Early in the development phase, TEAs of prototypes of training systems support and drive improvements of the design. As the system enters production and fielding, these evaluations provide quality control and help to identify any resource, personnel, or training issues across schoolhouses or units that may impact the effectiveness of the system. Finally, as the training system needs to be updated to maintain concurrency, leverage new technologies, and train for new tactics, techniques, and procedures, TEAs insure that those new capabilities deliver effective training. The Army recognizes the value of TEAs, and Army regulations require the regular assessment of training effectiveness of Training Aids, Devices, Simulations, and Simulators (TADSS). AR 350-52 requires that that all TADSS be evaluated to determine training effectiveness, cost effectiveness, cost trade-off, and cost savings. AR 350-1 and 350-38 require that TEAs be conducted on fielded systems to insure they continue to meet training needs throughout the lifecycle.

To the extent that the Army, and other services do conduct TEAs, the findings often rely heavily on 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-report. Although measuring Soldier learning through transfer to post-training exercises or an operational environment would provide a more valid measure of training effectiveness, the cost of obtaining these kinds of measures routinely is prohibitive. Costs are driven by a variety of

factors including the personnel, facilities, and resources needed for the data collection. Reducing those costs is key to enabling the Army to conduct TEAs throughout the lifecycle of its training systems.

To reduce the burden of collecting performance data from training systems, a methodology and supporting architecture are needed to derive objective metrics of Soldier performance within training technology systems. These metrics could be generated using the data these systems currently use 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 informing TEAs, there are several ways in which the development of these metrics could improve the execution and management of Soldier training. First of all, automated recording of trainee performance could serve to facilitate the diagnosis of performance problems and track progress over time. This would enable instructors to easily tailor remedial training to the individual Soldier. Automated recording of Soldier performance could also benefit training resource managers and researchers. Training resource managers could use these data to easily compare performance and resource use of multiple units over time. The training practices of units that consistently have better training outcomes could be implemented by all units to improve outcomes more broadly. Additionally, they could examine relationships between resource use and performance to determine the most cost-effective ways to deliver training.

Finally, intelligent or adaptive training systems could use these performance data to automatically recommend and deliver training to each Soldier based on their past performance. For example, within a simulator, an intelligent tutoring system could potentially provide feedback to each individual and even modify the training scenario to match the skill level of all students. Additionally, an intelligent tutoring system could recommend remedial training or drills to be completed outside of the simulator. To maximize the effectiveness of these data, they would be captured using a common standard specification, such as the Experience API (xAPI) so that they can be used to tailor training across multiple technology training systems.

In this paper, we describe a research effort designed to address the need for developing common metrics of Soldier performance using basic rifle marksmanship as a domain. In this effort, we will identify the objective metrics of performance that are primarily relevant to conducting TEA, but also address the needs of instructors, training resource managers, and other audiences. Unlike survey-based measures, these data will be gathered directly from Soldier performance across multiple marksmanship trainers and encoded in xAPI format. Using this specification will enable an objective evaluation of the effectiveness of the systems, tailoring training for individual Soldiers, and prescribing training across multiple platforms, among other uses.

BACKGROUND

Rifle Marksmanship has been one of the foundational combat skills of United States service members for almost 240 years. As such, both the Army and the Marine Corps have spent substantial resources in researching the most effective way of training marksmanship skills (Evans, Dyer, & Hagman, 2000). We chose marksmanship as a domain for this effort for several reasons. First, although the methods used to train this skill are well established and proven to be effective, the training and sustaining of this skill consumes a large number of resources including ammunition, ranges, and the time of Soldiers and their skilled instructors. Secondly, basic marksmanship is a skill that all Soldiers must master, regardless of their military operational specialty (MOS). Finally, while marksmanship may seem to be a straightforward procedural skill, it is in fact a deceptively complex task sensitive to a large number of cognitive and perceptual variables (Chung, et al, 2006).

As part of their initial military training, new recruits in the Army complete a course of instruction known as Basic Rifle Marksmanship (BRM) training. This training focuses on marksmanship fundamentals like body position, trigger squeeze, and sight picture and culminates with a qualification event in which trainees must hit a minimum number of timed targets at various distances. The biggest challenge for Drill Sergeants (DS's) conducting this training is that they have little time to provide one-on-one coaching for trainees. Additionally, they lack tools to quickly and reliably diagnose student problems, or to track student progress across periods of instruction. Because of this, most BRM training is a one-size-fits-all process. The priority of the DS's is to insure that every trainee meets the minimum

criteria with most resources being focused on the weakest shooters. For example, we have observed that approximately 20% of the trainees consume 50% of the ammo before they become proficient (Dyer et al, 2011). To put this in perspective, this means that of five trainees, half of the ammo is used to train four of them and the other half is used to train the fifth.

While marksmanship skills may seem very straightforward to train, the ability for a Warfighter to consistently hit a target at the distances required of Soldiers and Marines is actually quite remarkable. Soldiers are required to hit a target at a distance of 300 yards, or three football fields. This is challenging because even minute deviations from the center line of the rifle will result in large variations in where the bullet lands. Furthermore, there are myriad factors that can cause these variations. The Warfighter must learn to account for uncontrollable factors such as wind, temperature, gravity, and ammunition ballistics. Additionally the shooter must account for factors related to their own performance, including their breath, flinching, and wobbling due to fatigue. Even an increase in heart rate can affect shooter performance. Accurate marksmanship performance is a complex interaction of all these, and other factors (Chung et al, 2006). In order to master marksmanship, Warfighters must learn to master physical, cognitive, and emotional states as well as remember tactics, techniques, and procedures.

Marksmanship skills have been trained using several different technology systems since the 1980s, including Superdart, Weaponeer, the Laser Marksmanship Training System (LMTS) and the Army Research Institute's Multipurpose Arcade Combat Simulator (MACS). Currently, Soldiers use a simulator called the Engagement Skills Trainer (EST), while Marines use the Indoor Simulated Marksmanship Trainer (ISMT). These systems currently collect shot placement and aim-trace (the location of the aim-point for a few seconds before and after the trigger is pulled) for every shot from every Warfighter. Additionally, some live fire ranges have the ability to automatically detect target hits and others have sensors to determine the fall of the round (i.e., Location of Hits and Misses or LOMAH). Low-cost technologies are also available to automatically collect trigger-squeeze pressure, rifle movement, respiration, heart-rate, and many other measures from the shooter and rifle although these are not currently used during BRM training.

These marksmanship systems provide the benefit of training marksmanship safely while minimizing the costs associated with live range training. They vary in terms of cost, throughput, and fidelity. While the Army does an internal analysis of these systems prior to their deployment, a systematic comparison of their training effectiveness has not been completed. That said, the training effectiveness of several marksmanship systems has been evaluated over the years. These assessments vary widely in terms of methodology and participant sample. Importantly, there are no consistent metrics of performance used throughout these studies, which makes comparing findings across studies difficult. Table 1 presents a sampling of recent TEAs of marksmanship systems along with the metrics associated with their design. The table reports findings of published research with these systems. The current research will leverage additional marksmanship trainers, including several for which metrics have not been empirically investigated.

There is little disagreement that the ability of a Soldier to hit a target with a bullet should be a key metric of marksmanship performance. However, how that ability is operationalized varies across these and other studies. Measures of accuracy vary from the percentage and number of target hits to group size to the number of rounds to the first hit. Additionally, these studies vary in terms of the number of participants (from 18 to 386), the experience level of the participants (DoD Civilians with no marksmanship experience to experienced Infantrymen), and the control treatment (no training to live range training).

Table 1. Metrics of Marksmanship Performance in TEA Studies

Authors	System Evaluated	Metrics
Torre, et al. (1987)	Artificial Intelligence Direct Fire Research Test Bed	Rounds to zero; Group diameter; Hits; Proportion targets engaged; Rounds fired; Rounds to first hit; Time to fire
Eisley, et al. (1990)	Squad Engagement Training System (SETS)	Bullet strike efficiency (bullet holes/rounds fired); Targets hit efficiency (percentage hits); Tactical knowledge
Hagman (2000)	Laser Marksmanship Training System (LMTS)	Rounds fired; Trainees firing to standard; Known-distance hits; Qualification scores
Yates (2004)	Indoor Simulated Marksmanship Trainer (ISMT)	Qualification scores; Group diameter
Scribner, et al. (2007)	Dismounted Infantry Survivability and Lethality Testbed (DISALT)	Percentage of hits; Reaction time; Radial error from center of mass; Stress and Workload
Jenson & Woodson (2012)	ISMT	Mean point of impact
Robinson, et al. (2013)	ISMT	Subjective Survey
Getty (2014)	ISMT	Qualification scores; Group size; Mean point of impact; Center zeroing point length

In addition to the methodological inconsistencies in these studies, other variables likely contributed to their findings. For example, in Yates's (2004) evaluation of the ISMT, inclement weather impacted the test group but not the control group. Technical problems with the simulators resulted in significant data loss (Getty, 2014). When data were gathered, they were often highly variable (Torre et al. 1987). As a result, the findings from these studies are inconsistent, and tend to produce null results on the metrics that should be the most important, that is, the ability of a Soldier to accurately hit a target. Where differences are found, they tend to be in terms of efficiency metrics, suggesting that while simulation-based marksmanship training may not produce gains in Soldier accuracy or qualification scores, it does enable the same level of proficiency to be reached quickly and less expensively, which in itself demonstrates the value of these systems.

A step toward improving the methodology for conducting TEA of marksmanship systems would be to standardize the metrics used to evaluate Soldier performance. These metrics could be used to evaluate performance throughout the Soldier's career in current and future training systems. In addition to facilitating TEA, these metrics could be used to adapt training to the individual Soldier across systems, manage training resources, and inform researchers, among other uses.

DEVELOPING METRICS FOR INTEROPERABILITY

The Advanced Distributed Learning (ADL) Initiative, of the Office of the Under Secretary of Defense for Personnel and Readiness (OUSD P&R) is furthering the goals beyond interoperability towards learning ecosystems that provide a rich environment for connected training and learning. Moving beyond the Sharable Content Object Reference Model (SCORM) (Advanced Distributed Learning 2015a), ADL is working to build a foundation for Personal Assistants for Learning (PAL) through the Training and Learning Architecture (TLA) (Advanced Distributed Learning, 2015b). The TLA aims to allow the tracking of learning experiences with rich contextual data far beyond simply results and scores.   The TLA work is focused on extending the future support of interoperability of learning systems by making that data available for use in systems along the continuum of learning that a user experiences.

Under the TLA, the Experience API, or xAPI, enables tracking across platforms to capture data about learners and learning experiences (Advanced Distributed Learning 2015c). Whereas SCORM typically tracked browser-based e-learning, the xAPI has the capability to track data across simulation, virtual worlds, mobile applications, games and even data tracking from live training.

The xAPI defines a means to describe data and allows statements of experience to be created and stored in a simple storage mechanism called a Learning Record Store (LRS) (Advanced Distributed Learning, 2015d). The format of an xAPI statement is based on Activity Streams in the format of: <Actor, Verb, Object> or "I did this" (Activity Streams, 2015).

While the xAPI provides significant flexibility for encoding, best practices are emerging from communities of practice and research. Previous work by the Army Research Laboratory (ARL) on Interoperable Performance Assessment (IPA) focuses on uniformly defining and describing learning experiences along a continuum as a means to support adaptation and data analytics (Poeppelman et al., 2013). The IPA effort leverages the constructs of the Human Performance Measurement Language (HPML) (Stacy et al., 2006) and best practices to encode xAPI statements (Hruska et al 2013). The IPA research works primarily toward the goal of defining uniform performance measures in simulation and providing summative assessments towards these measures from multiple sources. Additional efforts by ARL to on IPA using small group and team data also indicate the potential of such approaches to adapt and even drive team formation (Hruska et al., 2014).

CURRENT RESEARCH

This effort will conduct the foundational research needed to provide guidance for how to develop effective metrics for marksmanship TEA, how to represent those metrics in xAPI format, how to design and develop an architecture to support the management analysis and interpretation of those metrics, and an interface so that instructors can easily access the data in meaningful ways. Additionally, the architecture would support the management of data for longer range studies.

To complete this research, we will leverage a number of live and virtual training systems already in use by the Army, Marine Corps, and other agencies in the Department of Defense that are available to automatically collect performance measures. For simulation based training, we will use the Army's Engagement Skills Trainer (EST), and the Marine Corps' ISMT. For live training, we will use the Marine Corps' Modular Advanced Technologies Marksmanship Proficiency (MAT-MP) system, which uses a set of sensors that are attached to a Soldier's service rifle to collect shooter performance data during live-fire training. Our goal is to develop metrics that can be used across these different marksmanship trainers to evaluate Soldier performance. While our focus is primarily on metrics for TEA, we also plan to address the needs of several groups who could benefit from this research, including instructors, resource managers, and researchers.

User Needs Analysis

Our first task will focus on developing use cases that define the method for measuring marksmanship performance and contextual information in the trainers described above. These will be developed based on training observations at FT Benning, interviews with Subject Matter Experts (SMEs), and a review of the literature to date. The process will focus on user needs from three communities: marksmanship instructors, training resource managers, and researchers. The outcome of the task will be a user needs analysis that can be used to develop metrics used by Army trainers and other personnel to evaluate Soldier marksmanship performance across multiple periods of instruction.

To fully understand user needs, we will conduct training observations at marksmanship training sites where the training technologies are being used. We will investigate where in the training cycle these technologies are used, how frequently they are used, and what the burden in terms of resources is. We will conduct interviews with Subject Matter Experts from each of our target user groups to identify which measures of performance are most important for their needs.

A key to being able to properly interpret performance measures is the inclusion of the right contextual data. A critical role for the user needs analysis is to identify contextual data that is necessary to know.   For example, if a

student fires at a target, we may record whether or not the target was hit. To understand the meaning of that hit or miss, we will need to know about any number of factors that might have contributed to that result. For example we might need to know: the range of the target, the shooter's firing position, if the shooter was wearing body armor, the shooter's height, the type of weapon used, the experience level of the shooter, the type of target, the ability of the instructor, the weather conditions, etc. Any and all of these factors can impact performance. To be able to use performance data to understand and correct problems, be they at the shooter level, instructor level, facility level, or course level, we have to be able to associate this contextual data with the performance measures we record.

Finally, the user needs analysis will help us to understand how to summarize or visualize these data for the different communities to use.   All of these communities will want to have tools to quickly and easily extract the information they need. The challenge here will be developing a user interface that is both intuitive yet also flexible enough to allow for customized reports or summaries for each community of users.

Metric Development

Based on the findings of the user needs analysis, metrics of marksmanship performance and contextual data will be developed for inclusion in the xAPI architecture. These metrics will be reviewed by subject matter experts (SMEs) from our targeted user communities to ensure they are appropriate and comprehensive. This review will take place during site visits or over telephone interviews.

In this process, we anticipate a number of challenges. The primary challenge will be identifying measures of performance and contextual data that are collected consistently across these systems without requiring further development. These systems are often used at different points throughout the marksmanship training process, and it is likely that they measure different outcomes. Another challenge involves developing metrics that are responsive to the needs of each of our target user groups. A key takeaway from our review of the marksmanship TEA literature is that these training systems collect vast amounts of data, which can be combined to develop many different metrics. While we aim to be responsive to the needs of several user groups, these users have different priorities. It is likely that they will need not only different metrics, but metrics on different levels of granularity. For example, an instructor may be satisfied with a Soldier's qualification score. However, a researcher may want more refined data to conduct more diagnostic analyses. Training systems have the potential to output vast amounts of data, and while one option for dealing with this challenge is to hold raw data in a repository and calculate metrics as the needs arise, ideally we would be able to anticipate both current and future needs to avoid the security and storage issues involved with that solution.

Architecture Development

A reference architecture will be developed for this research. The reference architecture will be composed of open source and government tools and will allow the collection and usage of the data. It will demonstrate the user needs and will provide a reference implementation of the long term technical architecture. The architecture will provide integration points with the identified marksmanship training systems. The architecture will leverage tools developed by ARL in the IPA research effort. Tools like the Soldier Performance Planner (SP2) and the xAPI encoding library, Pipeline, will be leveraged to collect and manage the data for the domain (Hruska et al, 2014). An architecture that allows both system and observer based measures will enabled through the research. The architecture will provide a baseline for future domain data collection.

Effectiveness Evaluation

Marksmanship is a skill which is developed across ten periods of BRM instruction in the Army. By automating performance measures across many of these periods, we will reduce the labor required to understand how performance in one period is both predicted by events in previous periods of training and predicts performance in subsequent periods. In this way, the effectiveness of a training intervention in any period can be easily determined. To test this capability we will look at training interventions during simulation training (EST) as well as live training.

While the focus of this research will be on training effectiveness assessment, candidate metrics will include those required by other audiences. For example, it will be possible to provide instructors with summaries of skill

progression by individual Soldiers. We will develop tools to quickly identify students in need of additional remedial training as well as to recommend specific training for those students. Additionally, tools will be developed for training resource managers to help track skill progression and resource use across units, facilities, etc. The tools for conducting training effectiveness assessments will be more sophisticated and complete than those built for these other audiences.

IMPORTANCE

This paper lays out a research plan for developing an architecture to utilize marksmanship performance measurement data to more effectively and efficiently deliver and manage training. The authors see this kind of architecture as a critical evolutionary step in the development of technology enabled training within the department of defense. Such an architecture will enable the military to move beyond training as a collection of separate but related events comprised of live, virtual, constructive, and gaming experiences to a truly integrated and adaptive enterprise.

The development of this architecture will support not only trainers and learners, but also improve training management and the acquisition and improvement of training technologies. By providing broad access to a rich human performance database, along with tools for analyzing and interpreting those data, it will be possible make rapid, data-driven decisions in these training, management, and acquisition domains.

REFERENCES

Activity Streams. (2014). *www.activitystrea.ms*. Retrieved 2014, from http://activitystrea.ms/

Advanced Distributed Learning. (2015a). *Sharable Content Object Reference Model*. Retrieved January 2015, from http://www.adlnet.gov/overview

Advanced Distributed Learning. (2015b). *Training and Learning Architecture (TLA): Experience API (xAPI)*. Retrieved from http://www.adlnet.gov/tla/experience-api

Advanced Distributed Learning. (2015c). *xAPI Specification*. Retrieved 2015, from Github: https://github.com/adlnet/xAPI-Spec

Advanced Distributed Learning. (2015d). *ADL Learning Record Store*. Retrieved 2015, from www.adlnet.gov: https://lrs.adlnet.gov/xapi/

Chung, G. K. W. K., Delacruz, G.C., de Vries, L.F., Kim, J., Bewley, W. L., de Souza e Silva, A. A. Sylvester, R. M., & Baker, E. L. (2006). Determinants of rifle marksmanship performance: predicting shooting performance with advanced distributed learning assessments. Los Angeles, CA: National Center for Research on Evaluation, Standards, and Student Testing.

Ciavarelli, A, Platte, W. L., & Powers, J. J. (2009). Teaching and assessing complex skills in simulation with application to rifle marksmanship training. Proceedings of the Interservice/Industry Training Simulation and Education Conference.

Dyer, J.L., Lipinski, J.J., Schaffer, P.S., Goodwin, G.A., James, D.R., & Dublac, M.D. (2011). Assessment of New Marksmanship Strategies in 2010. (ARI Research Report 1955). Fort Belvoir, VA: U.S. Army Research Institute for the Social and Behavioral Sciences. Unpublished Data.

Eisley, M., Hagman, J., Ashworth, R., & Viner, M. (1990). Training Effectiveness Evaluation of the Squad Engagement Training System (SETS) (Research Report 1562). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences.

Evans, K. L., & Osborne, A. D. (1985). A study of effectiveness of infantry systems: training effectiveness analysis, cost, and training effectiveness analysis and human factors in systems development and fielding (Research Note 85-38). Alexandria, VA: U.S. Army Research Institute for the Social and Behavioral Sciences.

Getty, T. J. (2014). A comparison of current naval marksmanship training versus simulation-based marksmanship training with the use of Indoor Simulated Marksmanship Trainer (ISMT) (Masters' Thesis). Monterey, CA: Naval Postgraduate School.

Hagman, J. D. (2000). Basic rifle marksmanship training with the laser marksmanship training system (Research Report 1761). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Hruska, M., Poeppelman, T. R., Dewey, M., Paonessa, G., Paonessa, M., Nucci, C., & Ayers, J. (2013). Interoperable performance tracking to support tailored learning (Final Report). Orlando, FL: U.S. Army Research Laboratory Human Research and Effectiveness Directorate.

Hruska, M., Long, R., Amburn, C., Kilcullen, T., & Poeppelman, T. (2014). Experience API and team evaluation: evolving interoperable performance assessment. Proceedings of the Interservice/Industry Training, Simulation & Education Conference.

Hruska, M., Long, R., & Amburn, C. (September, 2014). Human Performance Interoperability via xAPI: Current Military Outreach Efforts. Fall Simulation Interoperability Workshop, 14F-SIW-035, Orlando, FL.

James, D. R., & Dyer, J. L. (2011). Rifle marksmanship diagnostic and training guide (Research Product 2011-07). Fort Belvoir, VA: U.S. Army Research Institute for the Social and Behavioral Sciences.

Jensen, T. & Woodsen, J. (20012). Enable marksmanship training transfer study: The use of indoor simulated marksmanship trainers to train for live fire. (Master's thesis) Monterey, CA: Naval Postgraduate School.

Poeppelman, T. R., Hruska, M., Ayers, J., Long, R., Amburn, C., & Bink, M. (2013). Interoperable performance assessment using the Experience API. Proceedings of the Interservice/Industry Training, Simulation and Education Conference.

Scribner, D. R., Wiley, P. H., & Harper, W. H. (2007). A comparison of live and simulated fire soldier shooting performance (ARL-TR-4234). Aberdeen Proving Ground MD: U.S. Army Research Laboratory.

Stacy, W., Ayers, J., Freeman, J., & Haimson, C. (2006). Representing human performance with human performance measurement language. Washington, DC. Aptima, Inc.

Torre, J. P., Maxey, J. L., & Piper, A. S. (1987). Live fire and simulator marksmanship performance with M16A1 rifle: Study1. A validation of the artificial intelligence direct fire weapon research test bed. United States Army, U.S. Army Research Project Manager for Training Devices. U.S. Army.

Williams, H.; Robinson, E.; Kirkendall, C. (2013). Development of a subjective evaluation tool for assessing marksmanship training effectiveness (NAMRU-D Report 13-2). Dayton, OH: Naval Medical Research Unit Dayton.

Worden, B. P. (2012). Increasing realism in virtual marksmanship simulators (Master's Thesis). Naval Postgraduate School, Monterey, CA.

Yates, W. (2004). A training transfer study of the Indoor Simulated Marksmanship Trainer (Master's Thesis). Naval Postgraduate School, Monterey, CA.