Interoperable Performance Assessment using the Experience API

Tiffany R. Poeppelman, Jeanine Ayers
Aptima, Inc.
Fairborn, OH
tpoeppelman@aptima.com, jayers@aptima.com

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

Rodney Long, Charles Amburn
ARL STTC HRED
Orlando, FL
rodney.long@us.army.mil, charles.amburn@us.army.mil

Martin Bink
Army Research Institute
Ft. Benning, GA
martin.l.bink2.civ@mail.mil

ABSTRACT

Training technologies are evolving rapidly; yet, more efficient and effective use of resources and training time has not been significantly impacted by these advancements. Adaptive and tailored opportunities for learning represent a path ahead to larger efficiency (Durlach & Ray, 2011), but without interoperable tracking and assessment from multiple training systems, improved solutions are potentially cost prohibitive. By leveraging analytics and metrics of performance-based activity data from a variety of sources, organizations can provide the right support to unlock potential efficiencies. A number of efforts are underway at the Department of Defense (DoD), to develop adaptive, learner-centric systems that drive learning and performance forward. Army efforts, driven by Army Learning Concept 2015 (Department of the Army, 2011), such as the Soldier Centered Army Learning Environment (SCALE), and the Generalized Intelligent Framework for Tutoring (GIFT) are working towards this end. Additionally, the Advanced Distributed Learning (ADL) Initiative is stewarding efforts, which focus on new approaches with standards and frameworks, such as the Training and Learning Architecture (TLA), and the Experience Application Programming Interface (xAPI). In concert, these efforts represent a path ahead to tailored and adaptive learning. In this paper, we will describe an approach that combines a subset of the previously described DoD learning and performance efforts to create an interoperable method for tracking individual and group performance across mobile interactions, virtual worlds, real world activities, games, and other sources. This method will show that data about informal and experiential learning can be integrated with traditional formal sources and other organization systems easily. The resulting image will highlight a better understanding of the connection between learning and performance data.

ABOUT THE AUTHORS

Tiffany Poeppelman is an Industrial-Organizational Scientist at Aptima with experience in both the commercial and government sector, specifically regarding cutting-edge adaptive training technologies, training development, and applied research. Ms. Poeppelman is a seasoned training specialist who designs, develops, and executes training within various training environments, such as Live-Virtual-Constructive (LVC), computer-based, instructor-led, simulation-based, and game-based. She also conducts functional research to create innovative adaptive training solutions that adjust training experiences via human performance algorithms, scenario engineering, scaffolding, and other methods of intelligent adaptation. Ms. Poeppelman received a M.S. in Industrial and Organizational Psychology from Northern Kentucky University and a B.S. in Psychology from Bowling Green State University.

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, MD. He is currently the President/CEO 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 API. He holds a B.S. 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).

Jeanine Ayers is the Director of Technology Strategy at Aptima. She leads teams of software engineers in development efforts focused on architecting and building data-driven software applications in support of multi-agent systems, computer based training systems, and performance measurement and assessment systems. Ms. Ayers has extensive experience with High Level Architecture (HLA) and Distributed Interactive Simulation (DIS) simulators in DoD environments, including work for the Army, Navy, and Air Force. She serves on the TRAIN SISO committee which promotes the use of simulations to support the acquisition of knowledge and skills through education, training, and performance support. Ms. Ayers has an M.B.A. from Boston University and a B.S. from Carnegie Mellon University in Industrial Management/Information Systems.

Rodney Long is a Science and Technology Manager at the Army Research Laboratory, Human Research and Engineering Directorate, Simulation and Training Technology Center (STTC) in Orlando, Florida. He is currently the STTC project lead for SCALE, conducting research in support of the new Army Learning Model. Mr. Long is also the STTC project lead for the Joint and Coalition Training Rehearsal and Exercise Research (JCTRER), supporting the use of LVC simulations for joint and coalition warfare. Mr. Long has a wide range of simulation and training experience that spans 25 years in the DoD and has a Bachelor's Degree in Computer Engineering from the University of South Carolina and Master's degree in Industrial Engineering from the University of Central Florida.

Chuck Amburn is an Instructional Systems Specialist for the US Army Research Laboratory, Human Research and Engineering Directorate, Simulation Training and Technology Center in Orlando, FL. After obtaining both a Film degree and a Master's degree in Instructional Systems Design from the University of Central Florida, he began his DoD 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 of Simulation, Training and Instrumentation (PEOSTRI), Orlando, Fl. Since 2011, Mr. Amburn is also the Senior Instructional Systems Specialist for the SCALE research program.

Martin Bink is a Senior Research Psychologist at the U.S. Army Research Institute for the Social and Behavioral Sciences – Fort Benning. Dr. Bink’s research interests are in human learning, memory, and cognition especially as applied to education and training, and his current research focuses on digital-skills training, aviation-skills training, simulation training, and adaptive training methods. Dr. Bink holds a Ph. D. in Cognitive Psychology from The University of Georgia.

Interoperable Performance Assessment using the Experience API

Tiffany R. Poeppelman, Jeanine Ayers
Aptima, Inc.
Fairborn, OH
tpoeppelman@aptima.com, jayers@aptima.com

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

Rodney Long, Charles Amburn
ARL STTC HRED
Orlando, FL
rodney.long@us.army.mil, charles.amburn@us.army.mil

Martin Bink
Army Research Institute
Ft. Benning, GA
martin.l.bink2.civ@mail.mil

INTRODUCTION

The Need for Interoperable Performance Assessment

Organizations in the U.S. spent approximately $156.2 billion on employee learning and development in 2011 (ASTD, 2012). This investment is increasingly being spent on new training technology and methods of delivery such as mobile learning, virtual collaborative workspaces, and distributed simulations (ASTD, 2012). As the complexity of learning environments grow, so too does the need for more effective measurement of learner performance and state. For example, simulations are designed to provide learners with realistic *practice*. Simulator-based training, on the other hand, must provide *deliberate practice with feedback* (Stacy et al., 2006). Without an assessment of learner proficiency, it is impossible to measure whether a training program is effective.

Additionally, new opportunities for performance assessment and data collection continue to increase. A flux of more available data is occurring with the release of new methods for capturing learning data, such as Experience Application Programming Interface or API (xAPI) and learning and performance data-centric systems like a Learning Record Store (LRS). Yet, significant performance improvements have not been demonstrated. Others are working on solutions to this problem by leveraging adaptive training methods and intelligence tutoring frameworks. The Army's Generalized Intelligent Framework for Tutoring (GIFT) is working to inform the trainee of their performance during and after a training event.

Lastly, delivery of learning content ranges in methods from instructor-led and self-paced to continually expanding technology options. As organizations develop an appreciation for technology-based training and the demands for personalized learning continue to increase, it is imperative that researchers begin to connect these training technologies, data management solutions, and intelligent tutoring frameworks to maximize training efficiency. The current effort described in this paper highlights a new architecture and approach for capturing the right data and connecting cutting-edge frameworks and defined systems as well as enabling interoperable performance assessment in real-time. The development of tools and systems that are consumers and producers of adaptation data, like the Soldier Performance Planner (SP2) described in this paper, will not only inform future organizational best practices, but will also guide forward thinking towards system connectivity in future learning ecosystems. An ability to continually track and assess the performance of an individual and ultimately a team of individuals *independent of the delivery platform* is required to realize truly effective and efficient training.

State of Current Learning Environments

Currently, computing power and connectivity are evolving rapidly; therefore, organizations are not only choosing a combination of training methods (ASTD, 2012), but are also relying on technology-based solutions as a viable alternative. Training technology offers a wide range of benefits, including reduction in cost and training time increase in pedagogical capabilities, diversity, accessibility, and beyond (Bell & Kozlowski, 2007). While types of technology range from basic to advanced, each comes with their own capabilities, which have evolved training as we know it (Figure 1).

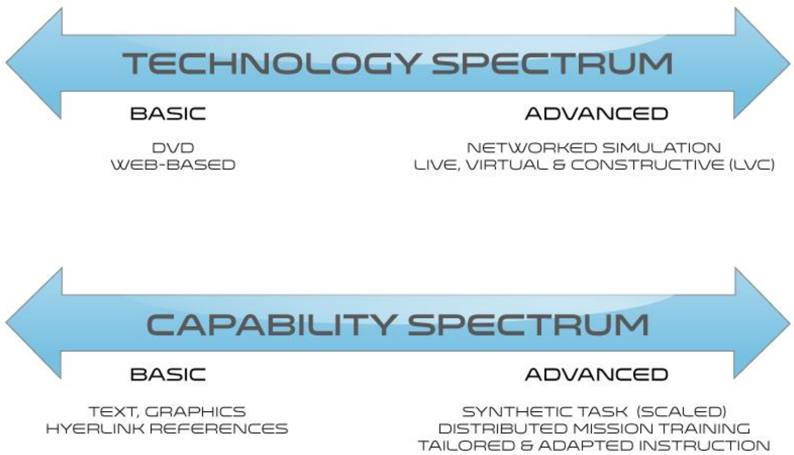

Figure 1. Training technology and capabilities spectrum

Organizations that are paving the way with cutting-edge training technology include the U.S. Army and U.S. Air Force who have invested heavily in the development of networked training devices that operate in shared virtual environments (human-in-the-loop simulators) in order to prepare units for battle. These environments extend naturally to Live-Virtual-Constructive (LVC) training opportunities, where live platforms are integrated with virtual and constructive systems to form an integrated training solution. With this increasing migration towards integrated LVC training, however, there is an increasing need to understand human performance both within and across systems and environments. LVC training provides a rich opportunity to assess performance, but the current problem is that performance measurement data is often incompatible across systems. The current methods for tracking performance in LVC are often stove-piped and rely on either human observer/controllers or pods which record and store the platform data during the exercise. In addition, a deficiency in a common format or communication protocol to collect and aggregate data exists. As a result, it is nearly impossible to effectively and efficiently generate, identify, extract, and track LVC performance data, thus rendering training less effective than the actual potential. Without an assessment of the state of the trainees, it is impossible to adapt and tailor learning experiences accurately.

Efforts that drive learning systems towards tailored and adaptive capabilities are important. A number of efforts are underway at the Department of Defense (DoD), to develop adaptive, learner-centric systems. The efforts are aimed at delivering lifelong, persistent, and ubiquitous learning opportunities to users where each advancement will provide incremental value in understanding the connection between learning and performance data. These efforts require that informal and experiential learning be integrated with traditional formal sources and other management systems to produce the desired capability. A key component to the success of adaptive learning is an interoperable method for tracking individual and group performance across the variety of mobile interactions, virtual worlds, real-world activities, games, and other sources.

Army Learning Concept 2015

The U.S. Army is focusing on creating a more competitive environment for leaders and Soldiers through its new campaign for lifelong learning. With today's changing environment, the Army is moving outside of the schoolhouse learning approach for training and expanding their goals to include new training methods and forms of tracking performance over time (Department of the Army, 2011). The Army Learning Concept (ALC) 2015 (Department of the Army, 2011) introduces new concepts that have the potential to guide research and technical architecture. One major concept that was introduced as part of the vision for ALC 2015 is "relevant, tailored, and engaging learning experiences" as well as the need for adaptive learning capabilities for the current and future Army forces. Research and technical development in this area is pertinent to implementing 21st century soldier competencies, establishing a learner-centric 2015 learning environment, and developing the adaptive learning system infrastructure outlined in ALC 2015.

The foundation for adaptive and personalized training must start with the ability to measure and track performance of a learner in any training event and then use that information to produce the following items outlined in ALC 2015:

- Training that can be adapted for the Soldier's operational environment
- Training that adapts to the technical capabilities at the Soldier's point-of-instruction
- Training that adapts to the Soldier's experiences and skill level

In order to meet these new requirements, next generation training tools need to communicate in interoperable ways. Systems that aim to provide adaptive or tailored learning will need to leverage both data about learners and data about the content in continually evolving complexity, at more and more granular levels. In the next section, current Army efforts being leveraged for the current research to create an interoperable performance assessment prototype and open-source architecture will be described.

CURRENT ARMY EFFORTS

Soldier Centered Army Learning Environment (SCALE)

Developed by the Army Research Laboratory (ARL) and Army Research Institute (ARI), the SCALE (Mangold, Beauchat, Long, & Amburn, 2012) is intended to provide a data-driven architecture to support tailored training and education across multiple hardware platforms (personal computer and mobile devices), using mobile applications, virtual classrooms and virtual worlds. Of the ALC 2015's 25 required capabilities, SCALE addresses the following relevant categories:

- Learning environment architectures
- Adaptive learning and intelligent tutoring
- Advanced networking
- Learning management
- Knowledge engineering

The aim of SCALE is to be modular and web service-based in order to facilitate the integration of new technologies into the broader SCALE architecture, and to potentially allow the integration of SCALE into existing and new technologies. This scalability and extendibility will serve itself well with other technologies that are being built for ALC 2015. A key requirement of the SCALE portal is the incorporation of trainee performance assessment data that can be used by other SCALE components and 3rd party training systems. Informing this SCALE component is one of the main focuses of this research effort.

Generalized Intelligent Framework for Tutoring (GIFT)

One potential consumer of performance and assessment data from SCALE is the GIFT. Developed by ARL, GIFT is a modular tutoring system framework to support the instantiation of adaptive tutoring capabilities. GIFT aims to research and prototype a computer-based tutoring framework to evaluate adaptive tutoring concepts, models, authoring capabilities, and instructional strategies (Sottilare, Goldberg, Brawner, Holden, 2012). The effort aims to do this across various populations, training tasks, and conditions, thus enabling summative and formative evaluations. GIFT services provide authoring of Computer-based Training Systems (CBTS) and CBTS components, tools and methods; management of instructional processes using best pedagogical practices based on the behaviors of expert human tutors; and an assessment methodology to evaluate the effectiveness of CBTS and CBTS components, tools, and methods. GIFT provides services for learners, instructional system designers, expert behavior modelers, training system developers, trainers, and researchers.

GIFT's infrastructure provides generic tutoring or remediation strategies to integrating systems based on learner performance. Currently, ARL is in the process of integrating GIFT as a pedagogical feedback mechanism within the SCALE environment. The addition of GIFT into SCALE will allow for real-time intelligent tutoring and remediation as learners progress through future training.

Aviation Collective Performance Assessment Toolset (AC-PAT)

Aviation Collective Performance Assessment Toolset (AC-PAT) is a suite of performance measurement tools for aviators currently under development for ARI. AC-PAT is composed of two measurement tools. The first tool is a performance measurement application that (1) captures simulation data from a distributed virtual training environment and (2) automatically calculates system-based measures of aviation collective performance. The second tool is a touch screen tablet application that enables Observers/Controllers or Instructor Pilots to capture behaviorally-anchored assessments of pilot performance. Together, these two tools provide a comprehensive view of performance that enables training professionals to understand, not just how well a trainee performs, but why and where they could improve.

However, these two measurement modalities (i.e., observer-based and system-based) do not currently integrate data or feedback. That is, each modality can be used to measure simulation-training performance, but the data from each modality cannot be aggregated to provide multi-modal feedback. Moreover, there is no mechanism to store the data from either modality in a way that tracks training performance across a unit or an individual. So, even though the current ARI AC-PAT provides effective feedback for individual training events, a more complete measurement system is required to provide (a) multi-modal measurement and feedback, (b) individual training progression across time, and (c) training aggregation across individuals in a unit. A current ongoing effort that addresses some of these issues is the Advanced Distributed Learning (ADL) Initiative which will provide the capability to gather new types of performance data that was not previously being tracked on a large scale.

Experience API

In 1999, the ADL Initiative, currently of the Office of the Under Secretary of Defense Personnel and Readiness (OUSD P&R), set out to modernize learning and training in the DoD. As a background, one of ADL's initial goals was to "establish guidelines on the use of standards and provide a mechanism to assist DoD and other Federal agencies in large-scale development, implementation, and assessment of interoperable and reusable learning systems" (Training Industry, 2013). As a result of the creation of ADL, the Sharable Content Object Reference Model (SCORM) emerged (Advanced Distributed Learning, 2013).

The ADL Initiative is currently working to extend the future support of interoperability of learning systems through a Training and Learning Architecture (TLA) (Advanced Distributed Learning, 2013). The ADL initiative is working with communities of vendors and users to develop specifications to support learning technology beyond the SCORM. The first project of the TLA effort is known as the Experience API or xAPI (Advanced Distributed Learning, 2013). The xAPI specification (Advanced Distributed Learning, 2013), 1.0.0 as of April 26, 2013, defines an API to track data about learning experiences. An API defines a protocol intended to be used as an interface by software components to communicate with each other (Advanced Distributed Learning, 2013). The xAPI defines a method to capture data about the interaction between a learner and a learning experience.

Unlike SCORM, the xAPI allows tracking outside of an LMS to capture data about digital and non-digital learning experiences. The xAPI allows the interoperable tracking of user experiences in newer content types like serious games, mobile applications, simulations, virtual world, augmented reality, and observer-based measures. As data is collected via the xAPI, it can also be exposed so that it can be used for assessment after the fact, statistical analysis, data mining, custom reports, and data sharing between systems.

The xAPI is based upon an open format specification for activity stream protocols, which are used to syndicate activities (Activity Streams, 2013). The xAPI specification defines a format using JavaScript Object Notation (JSON) to create statements that allow the capture of learning experiences. JSON is a text-based open standard designed for human-readable data interchange. The format of a statement is <Actor, Verb, Object> or "I did this."

Producing effective adaptive and tailored learning experiences requires data. The xAPI provides a standardized capability to capture significant amounts of granular data across devices. The xAPI allows the capture of data about the user, content, and context associated with an experience. Additionally, the xAPI can capture extended data about outcomes and results. This standardized capture was not possible through a uniform method and has not been previously available across platforms.

The following (Figure 2) is an example of an xAPI statement in JSON.


```
{
   "actor": {
      "name": "Mike Smith",
      "mbox": "mailto:mike.smith@us.army.mil"
   },
   "verb": {
      "id": "http://adlnet.gov/expapi/verbs/experienced",
      "display": {"en-US": "experienced"}
      },
   "object": {
   "id": "http://example.com/activities/AH-64-Close-Air-Support",
      "definition": {
         "name": { "en-US": "AH-64-Close-Air-Support"}
    }}
```


Figure 2. xAPI statement examples

Statements can also be expanded to capture additional data about experiences. Outcomes and measures are also generally associated with many learning experiences. Accordingly, statements can also capture results of the learning experiences. The xAPI format provides a flexible and interoperable way to capture key data about a user's learning experience. In this manner, it supports opportunities for next generation technologies to track and report across learning systems.

Handling Statements

The storage and retrieval of statements is critical to any solution that intends to leverage xAPI. While the xAPI defines the format for statements, it does not define how the technology will work to store them. The notion of a Learning Record Store, or LRS, is associated with the storage of statements (Advanced Distributed Learning, 2013). An LRS provides a single point to store statements and provides an interface to collect data about a user. An operationalized LRS can provide a profile of a learner's experiences and their corresponding data.

Human Performance Measurement Language (HPML)

Pervasive measurement and feedback is not only critical to creating adaptive soldiers, but is also a requirement for engineering an adaptive training system in accordance with the Army Learning Concept for 2015. The only way to have learner-centric approaches is to have accurate measurement systems to inform instructors, students, and adaptive training systems of the current learner state. These systems must provide reliable trend analysis information to instructors and curriculum developers as well. However, it is often difficult and time-consuming to construct and configure measurement and feedback data in a standard and reusable format that can be communicated, stored, and interpreted by various computer-based systems. The approach that will be described in this paper leverages Human Performance Measurement Language (HPML) a proven technology for representing theory-based frameworks for human performance in training (Stacy et al., 2006; Stacy, et al., 2005). HPML provides the ability to define, represent, and communicate trainee performance across various Army training programs. Specifically, separating theoretical constructs like measurements, assessments, instances of both, formalisms for describing context and its relationship to measurements and assessments, allows for the integration of simulator-based, observer-based, self-report, and physiological measures, and the separation of session data from measure computation.

HPML is defined as an XML schema designed to express performance measurement concepts in a format that is both machine and human readable. These concepts include team and individual performance criteria, measurement definitions that describe the data and calculations required to assess performance for a training entity or team, and contextual data used to trigger and or modify the assessment of performance. The following (Figure 3) is a small snippet of HPML that describes the requirements for measuring how a pilot will perform a certain event of interest in a simulated training mission.

```
<MeasurementGroup ID="WeaponEffect">
  <Property Name="MainMeasurement" Value="FINALWpn_Effects_on_Tgt"/>
  <Property Name="SupportingMeasure" Value="FiredWeapon"/>
  <Property Name="SupportingMeasure" Value="WeaponDetonated"/>
  <Property Name="SupportingMeasure" Value="TargetDestroyed"/>

  <Measurement ID="TargetDestroyed">
    <Parameter Name="DetonationResult" Type="Double"/>
    <ConstantValue Name="One" Type="Double" Value="1"/>
    <MeasurementComputation Operator="EQUAL">
      <ConstantRef Ref="One"/>
      <ParameterRef Ref="DetonationResult"/>
    </MeasurementComputation>
  </Measurement>

  <MeasurementComponent ID="DetonationResult">
      <Parameter Name="EntityID" Type="String"/>
      <DataSourceQuery From="Detonation" Type="INTERACTION" Select="DetonationResult"
      Where="FiringEntityID == %P%EntityID" DataSourceRef="DIS"/>
  </MeasurementComponent>

  <MeasurementTrigger ID="WeaponEffectTrigger">
        <TriggerStart Type="Measurement" Ref="WeaponFireTrigger" ResetOnTrigger="True"/>
        <TriggerStop Type="Measurement" Ref="FifteenSeconds"/>
  </MeasurementTrigger>
</MeasurementGroup>
```

Figure 3. Measurement definition snippet

The figure above illustrates a basic HPML example to measure the result of a pilot's weapon fire on a target. Three common elements of HPML are illustrated: *MeasurementGroup*, *Measurement*, and *MeasurementComponent*. The *MeasurementComponent* is the most granular element of a measure and defines an individual piece of data that is used by a Measurement. The Measurement then combines one to many *MeasurementComponents* with a computation for a given training entity or entities. The *MeasurementGroup* contains the highest level instructions and associations for combining measures and triggers to deliver an assessment of performance for a given training objective. In the example defined in Figure 3, the *MeasurementGroup* (line1) defines the measure called *WeaponEffect*. This is the high-level measurement definition that associates individual measures required to calculate if the aircrew was able to destroy the target (note that the above definition has been abbreviated for illustration purposes). To begin, the *WeaponEffect* measure is composed of four different sub-measures, *FINALWpn_Effects_on_Tgt* (line 2), *FiredWeapon* (line 3), *WeaponDetonated* (line 4), and *TargetDestroyed* (line 5). Each of these sub-measures references a more specific measurement definition in a different location in the file. The *Measurement* called *TargetDestroyed* measures whether a detonation event destroyed a target. *TargetDestroyed* is composed of a *MeasurementComponent* that defines how to collect the required piece of individual data from a specific simulator environment to determine if a target is destroyed or not. Finally, the *Measurement* called *WeaponeFireTrigger* is the event that will start the calculation of the *WeaponEffect* measure for a given entity of interest.

HPML provides a way to describe performance in complex ways for individuals. Data described and captured using HPML allows performance to be understood in a real-time situation or after the fact. It also links performance measures to their associated training objectives, task conditions, and standards. HPML can handle the differences in collecting performance across LVC environments. Performance can also be described using HPML for groups of trainees.

A new approach to combining a number of DoD learning and performance efforts with xAPI and HPML to create an interoperable method for tracking individual and group performance across mobile interactions, virtual worlds, real world activities, games, and other sources will be described below.

THE APPROACH

The current approach of this development effort is to use the xAPI to capture performance data coded as HPML in a distributed training environment in a way that can be leveraged by other systems, like SCALE and GIFT, within the Army environment. The current method concentrates specifically on capturing data about context and performance within a distributed simulation training event to answer the questions like (1) What do the trainees learn during the event? and (2) Why did their performance increase or decrease?

As described above, HPML descriptions can be used to encode and extend xAPI statements. Capturing data from experiences using xAPI statements with additional description from HPML provides a new leverage point to support adaptive and tailored learning. Statements in this way can be encoded to contain data about the performance context in which they were captured. Data from experiences where both learning and performance occur can also be captured. Experience API data from these experiences can be used to customize subsequent learning experiences. If subsequent learning experiences can be customized based upon a user's performance within learning experiences, a potential exists to optimize or compress learning and the associated cost.

Each of the previously mentioned efforts at the Army and ADL will prove to be valuable to ensuring future training tools communicate in the ways necessary to support adaptive learning. In addition to the SCALE architecture and TLA model, the xAPI affords an interoperable mechanism to track users in a wide variety of environments. While the xAPI allows tracking of activity, very few researchers have begun to determine activities that can be captured in the context of performance.

The current research effort, which has begun to think in this way, will leverage data captured in an LRS, which has performance context encoded to inform systems to adapt or tailor learning experiences throughout Army systems. Data from the LRS can ultimately be used in automated ways to define parameters, inputs, and actions associated with adaptive learning, mitigations, and other interventions on an ongoing basis.

The following diagram (Figure 4) illustrates the overall system approach:

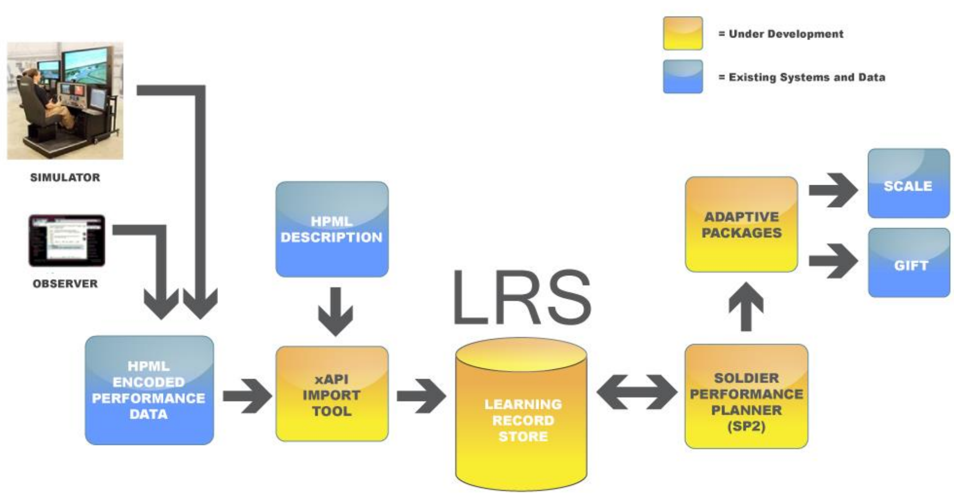

Figure 4. Interoperable performance assessment architecture

As mentioned, the xAPI uses vocabulary, activities, extensions, and additional encoding using HPML to form a basis to track user experience data. Because the tracked experience data contains descriptions of the experience (from xAPI statements) and data about the performance context (from HPML descriptions encoded in xAPI statements), researchers now have the ability to understand performance in a more real-time fashion. By leveraging the descriptions of performance from HPML and the encoding of xAPI statements with HPML data, we can view data stored in an LRS through the definitions of HPML. In this manner, one can view performance measures and their corresponding data from the LRS. Viewing can be done in real-time or after completion of an activity.

The current research effort is developing the SP2 which is an open-source prototype that allows the user to interface with the data in the LRS to evaluate performance measures which indicate proficiencies and deficiencies for an overall event or along a timeline. The SP2 can not only define deficiencies, but prepare a package or message to be sent to another Army system with information to adapt or tailor future learning. In this manner, HPML encoded xAPI statements are stored in an LRS, retrieved and analyzed by the SP2, and packaged to inform a different system with data to support adaptation.

While this approach defines a method for performance measures around individual learners, SP2 will be extended to track group performance. The same approach for performance measures will be undertaken with *performance indicators* that describe aggregate targets and measures of individual performance.

Current Army Use Case

The following example demonstrates the type of performance context that can be captured and used to understand the experience a trainee had during a real Army exercise. Based on a current effort with ARI, effective performance measures were collected for a variety of training techniques (e.g., observer-based, self-report, and system-based) through the COmpetency-based Measures for Performance ASsessment Systems (COMPASS) (Garrity et al., 2004) methodology that combines experiential knowledge of subject matter experts (SMEs) with established psychometric practices. The measures were represented in HPML and instantiated within AC-PAT software developed for ARI. The measurement software was then used to measure performance of Attack Weapons Teams (AWT) during a pre-deployment exercise.

The training event was an *Air to Ground Integration* for a *Collective Training Challenge* (associated with *Mission Essential Task List item 01-2-5183- Perform Tactical Air Movement Operations*). The training objective was "*Performed tactical air movement operations within the specified time constraints and according to commander's guidance*." The mission phase was *On-Station* and the performance indicator or observable behavior was *Employ Weapon System*.

In this effort, two different performance measures were captured. One of which was identified by the system and a second was measured by an observer. The observer captured whether the aircrew followed shot doctrine, which only an observer could assess, while the system measured whether the pilot delivered weapons in accordance with briefed parameters. In this particular example, the aircrew followed all shot doctrine but had an incorrect airspeed at weapon fire.

However, that is not where the data ends. The briefed parameters will detail which, if any of the many details a pilot must follow, were outside the bounds of acceptability. Thus, giving the pilot better feedback on how they performed and why they were evaluated in a particular manner could provide value. The data can also be used to deliver a more tailored learning experience in the next training opportunity. The following (Figure 5) is an example of contextual data that might be stored in extensions of the xAPI, using the example previously described.

```
"result": {
          "extensions": {
         "http://example.com/dodobjectiveregistry/learningobjective  ": {"name": " Perform Tactical Air Movement Operations ", "id" : "
http://example.com/dodobjectiveregistry/learningobjective/289209410,  "description": " Performed tactical  air movement operations within the
specified time constraints and according to commander's guidance", "measure": "Followed all shot doctrine", "assessment": "below average"},
"http://example.com/dodobjectiveregistry/learningobjective  ": {"name": "no friendlies in firing cone ", "id"
:"http://example.com/dodobjectiveregistry/learningobjective/289209410/12  , "description": " Within Tactical  Air Movement Operations, do not hit
friendly targets", "measure": "2 friendlies in firing cone", "assessment": "below average"},
},
"context": {
           "Training Objective": {
           "Measure": "Followed all shot doctrine",
           "Assessment": "below average"
          },
          "contextActivities" : {
          "category": [
                    {
                      "id": "http://www.example.com/meetings/categories/airtogroundintegrationevent",
                      "objectType": "Activity",
                      "definition": {
                        "name": {
                          "en": "Air to Ground  Integration event"
                        },
                        "description": {
                          "en": "Employ Weapon System"
                        },
                        "type": "http://example.com/expapi/activities/trainingactivity"
                      }
                    }
                ],
          }
          "extensions": {
         "http://example.com/jmetl  ": {"name": " Perform Tactical Air Movement Operations ", "item" : "01-2-5183"}
      },

}
```


Figure 5. xAPI example using HPML data

SP2 Assessment

The method described above shows that data about informal and experiential learning can be integrated with traditional formal sources and other organization systems easily. The prototype system, SP2, is being validated in order to assess whether the technology is properly leveraging data from multiple systems as well as producing improved results for training effectiveness. Four key questions are being explored with the prototype solution: (1) Is the level of detail and context of performance measurement being captured by system- or observer-based tools using the xAPI and HPML at the appropriate level? This initial goal is to ensure the system can tailor training on specific skills and be used to accurately record and/or automate feedback and future content. (2) Can detailed performance data be used to adapt training to the Soldier's experience and skill level for a given mission? (3) Is the technology standard flexible enough to apply to the many components of SCALE, including GIFT and 3rd party training environments? and (4) Does the approach scale past the individual to include team performance data that is being collected across military end users? The questions being examined will be presented and results will be discussed.

Future Challenges

Assuming that activities can be captured at a meaningful level of granularity using tools such as SP2, AC-PAT, and others, problems associated with understanding the data as it relates to performance for individuals as well as quantifying, measuring, and interpreting group performance presents future challenges. First, tracking individuals within a course or across their entire careers presents different technical and design challenges. These challenges relate to encoding performance indicators in xAPI and potential data size and scale problems. Second, challenges for team performance typically associated with big data may also be encountered. Integrating both system- and observer-based measures is equally important to be understood. Future research and development efforts should take these data requirements and questions into account.

CONCLUSIONS

By leveraging analytics and metrics of performance-based activity data from a variety of sources, organizations can provide the right support to unlock potential efficiencies and increase learning and performance outcomes. The importance of incorporating valid and reliable performance assessment across environments and training venues cannot be overstated. It is the fundamental requirement in the development of advanced learning software that content and experiences adapt to the knowledge level of the trainee and progress at a rate that presents a proximal zone of development for the learner (Vygotsky, 1978). Particularly during LVC training, considerable care must be taken to develop secure, technology-enabled, integrated assessments tailored to content and expected outcomes. By using the described approach and tools in this paper, these challenges can be met by collecting trainee performance data during a training exercise to positively impact training in a variety of ways. First, the data can be consumed by real-time displays of trainee performance, reducing instructor workload and enabling more timely instruction and accurate feedback to trainees. Second, trainee performance data can be used to dynamically adapt training scenarios to support the evolving instructional needs of individual trainees (Dean et al., 2011). Finally, once training is complete, performance results can be captured and used to (1) provide diagnostic feedback in debriefing and post-mission readiness reports, (2) aid in the detection of performance deficiencies and providing targeted remediation, (3) adapt and tailor the next best training option, and (4) identify trends in performance to support the identification of critical training shortfalls that require modification of training curriculum or scenarios.

REFERENCES

Activity Streams. (2013). *www.activitystrea.ms*. Retrieved 2013, from http://activitystrea.ms/

Advanced Distributed Learning. (2013). *ADL Learning Record Store*. Retrieved 2013, from www.adlnet.gov: https://lrs.adlnet.gov/xapi/

Advanced Distributed Learning. (2013). *Sharable Content Object Reference Model*. Retrieved May 2013, from http://www.adlnet.gov/overview

Advanced Distributed Learning. (2013). *Training and Learning Architecture (TLA)*. Retrieved from Advanced Distributed Learning: http://www.adlnet.gov/capabilities/tla

Advanced Distributed Learning. (2013). *Training and Learning Architecture (TLA): Experience API (xAPI)*. Retrieved from http://www.adlnet.gov/tla/experience-api

Advanced Distributed Learning. (2013). *xAPI Specification*. Retrieved 2013, from Github: https://github.com/adlnet/xAPI-Spec

ASTD. (2012). *2012 State of the Industry Report.* Alexandria: ASTD Press.

Bell, B. S., & Kozlowski, S. W. (2007). Advances in Technology-Based Training. *Managing Human Resources in North America*, 27-42.

Dean, C. S. (2011). Item response theory adapts training to disparately skilled trainees. *Proceedings of the Interservice/Industry Training, Simulation, & Education Conference* . Orlando, FL.

Department of the Army. (2011). *The U.S. Army Learning Concepts for 2015.* Washington: Department of Defense.

Durlach, P., & Ray, J. (2011). *Designing Adaptive Instructional Environments.* Orlando: U.S. Army Research Institute.

Garrity, M. J., Entin, E. B., Morley, R. M., & MacMillan, J. (2004). *A streamlined scenario-based methodology for evaluating distributed performance.* AFRL Research Report AFRL-HE-AZ-TR-2004-0130. Mesa, AZ: United States Air Force Research Laboratory, Human Effectiveness Directorate.

Mangold, L. V. (n.d.). An Architecture for a Soldier-Centered Learning Environment. *Simulation Interoperability Workshop.*

Sottilare, R., Brawner, K., Goldberg, B., & Holden, H. (2012). *The Generalized Intelligent Framework for Tutoring (GIFT).* Retrieved from https://litelab.arl.army.mil/system/files/2012_07_Sottilare_IFest_GIFT%20brief%20plus%20panel%20briefs_0.pdf

Stacy, W. M. (2005). A language for rapidly creating performance measures in simulators. *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference.* Orlando, FL.

Stacy, W., Ayers, J., Freeman, J., & Haimson, C. (2006). Representing human performance with human performance measurement language. *Proceedings of the Fall Simulation Interoperability Workshop.* Orlando, FL.

Training Industry. (2013). *Advanced Distributed Learning Intiative*. Retrieved from www.trainingindustry.com: http://www.trainingindustry.com/taxonomy/a/advanced-distributed-learning-(adl)-initiative.aspx

Vygotsky, L. (1978). Interaction between Learning and Development. In *Mind in Society* (pp. 79-91). Cambridge, MA: Harvard University Press.