Managing Learning and Tracking Performance across Multiple Mission Sets

Eric Watz, Peter Neubauer, John Kegley

Aptima, Inc.

Dayton, OH

ewatz@aptima.com, pneubauer@aptima.com, jkegley@aptima.com

Winston "Wink" Bennett, Jr.

Air Force Research Laboratory – 711th Human Performance Wing

Dayton, OH

winston.bennett@us.af.mil

ABSTRACT

Live and simulation-based training systems are data-rich environments with a wealth of human-performance measurements. Knowledge Management Systems (KMS) that fuse multiple data sources to predict current proficiency and future training needs are essential to managing localized and longitudinal learning goals for individuals, teams, and teams of team. This is the foundation for adaptive proficiency-based training that improves training efficiency and effectiveness. A KMS can take many forms but fall into three distinct categories of technology and capability. This paper illustrates three primary categories of managing learning that apply to the majority of KMSs available today. Recent advances in training technologies along with lessons learned from implementation in the simulation and live operations domains have made it possible for the Air Force Research Laboratory (AFRL) engineering and research teams to execute advanced capture, processing, and visualization of training research data. This paper delves into some challenges that exist to measuring, tracking, warehousing, and tracking proficiencies across human performance data sets within the U.S. Air Force (USAF) tactical domains. These challenges include tailoring assessments by domain, strategies for handling the ever-increasing complexity of training in live, virtual, and constructive (LVC) environments, and the need for operationally-relevant and scientifically-based methodologies and practices when measuring and tracking training. This paper draws upon lessons learned from implementation of an enterprise Knowledge Management System and looks at the way forward in terms of performance measurement, data storage, and proficiency prediction.

ABOUT THE AUTHORS

Eric Watz is the Chief Engineer for Aptima, Inc.'s Product Engineering group. He provides expertise in the areas of human performance measurement, learning management systems, and integration of cognitive modeling architectures with distributed simulation and training systems. Mr. Watz leads the technical research and development for a number of Air Force Research Laboratory efforts that focus on developing new and innovative capabilities to train operational personnel in the tactical fast jet, remote piloted aircraft, and other domains. Mr. Watz holds an M.S. in Computer Information Systems from the University of Phoenix, a B.S. in Computer Science from Arizona State University, and a B.S. in Human Biology from the University of Wisconsin-Green Bay.

Peter Neubauer is a Principal Software Engineer with Aptima, Inc. specializing in distributed training, automated systems-based human performance assessment, learning management systems, and software architecture. He leads the Performance Evaluation and Tracking System (PETS), the U.S. Air Force's automated human performance assessment solution. Mr. Neubauer also contributes experience with pilot proficiency and training needs prediction and with data warehousing. Mr. Neubauer holds an M.S. in Embedded Systems and a B.S.E. in Computer Systems Engineering from Arizona State University.

Dr. John Kegley is a Research Scientist in Aptima, Inc.'s Learning & Training Systems Division. Currently, he is supporting the Air Force's Distributed Mission Operations (DMO) and Live, Virtual, and Constructive (LVC) efforts, where he is assisting with training research design and evaluation. Dr. Kegley holds a Ph.D. in Human Factors/Industrial Psychology from Wright State University, a M.A. in Industrial and Organizational Psychology from Western Kentucky University, and a B.A. in Psychology from the University of South Florida.

Dr. Winston "Wink" Bennett, Jr. is a Senior Research Psychologist for the Warfighter Readiness Research Division located at Wright Patterson AFB Ohio. He is an Air Force Research Laboratory Research Fellow and a Fellow of the Association for Psychological Science, the Society for Military Psychology and the Society for Industrial and Organizational Psychology. He is currently leading research developing methods to monitor and routinely assess individual and team competencies and performance across live and virtual environments and evaluating game-based approaches for training, work design, and job restructuring. He is an Associate Editor for the journal Military Psychology and serves as a contributing editor and/or as a reviewer for other professional journals. He received his Ph.D. in Industrial Organizational Psychology from Texas A&M University in 1995.

Managing Learning and Tracking Performance across Multiple Mission Sets

Eric Watz, Peter Neubauer, John Kegley

Aptima, Inc.
Dayton, OH
ewatz@aptima.com, pneubauer@aptima.com,
jkegley@aptima.com

Dr. Winston "Wink" Bennett, Jr.

Air Force Research Laboratory –
Dayton, OH
winston.bennett@us.af.mil

INTRODUCTION AND PROBLEM

The use of live and simulation-based training environments is commonplace among today's military personnel. These environments are often rich with data from one or more sources that can be leveraged and transformed into measurements and assessments, which in turn can be used by a knowledge management system to manage individual and team learning. The ability to measure and track performance and manage learning for individuals, teams, and teams of teams is critical to our ability to maintain readiness. To ensure that today's warfighters maintain their skills at the optimal level of readiness in the most efficient manner possible, it is necessary to manage data, manage learning content, track performance, and effectively warehouse all that information. Today's mission sets, training needs, and training capabilities are continually evolving over time. Therefore, systems that adapt to warfighters' needs and present the correct training opportunities, at the proper intervals and at the required level of cognitive difficulty, are needed to optimally maintain warfighters readiness.

A first step in the management of learning is to identify the key experiences and mission contexts that are required to refine and develop learners' expertise. Pre-training development requires an understanding of organizational/unit culture and training needs, establishing job-task requirements and competencies, and investigating individual- and group-level knowledge and skill gaps (Cheng & Ho, 2001; Burke & Hutchins, 2007). This finding accentuates the high priority that job and task analysis techniques, organizational studies, and business trend analysis play in designing and executing high utility instructional programs (Arthur, Bennett, Edens, & Bell, 2003; Salas, Kraiger, & Smith-Jentsch, 2012). Participants or organizations do not view training as valid if it was developed using inadequate information about actual learning requirements and how they aligned with business goals (Alliger, Tannenbaum, Bennett, & Traver, 1998).

Research has provided guidance in terms of factors that should be managed during the training phase as well. Facilitators effectively modeled desirable behaviors and provided practice opportunities with consistent and relevant feedback regarding errors and successful performance (Burke & Hutchins, 2007; Salas, Kraiger, & Smith-Jentsch, 2012). Trainees have been encouraged to think about and visualize how the instructional material could be extended and applied to the operational or on-the-job environment (Axtell, Maitlis, & Yearta, 1996). It was also critical to ensure that technology- or simulation-based learning modalities possessed high psychological fidelity, supported structured user and instructor control, and enabled complementary data fusion to highlight areas for additional knowledge or skill reinforcement (Schreiber & Bennett, 2006; Salas, Kraiger, & Smith-Jentsch, 2012; Bennett, Tripp, Berlingo, & Watz, 2016).

BASIS FOR PROFICIENCY-BASED LEARNING MANAGEMENT AND TRACKING SYSTEM

Knowledge management systems (KMS) are "systems that help instructors provide their students with learning materials and activities while tracking participation and progress through data systems and assessments" (Falvo & Johnson, 2007, p. 40). Many different domains employ KMS platforms, including: K-12 education, undergraduate and graduate programs at colleges and universities, commercial and non-profit organizations, and across all the United States Department of Defense (DoD) agencies (Davenport & Prusak, 2000; Falvo & Johnson, 2007; Kamel, 2007; Watson & Watson, 2007; Goodwin, Johnson, Sottilare, Brawner, Sinatra, & Graesser, 2015; Bennett, Berlingo, Tripp, & Watz, 2016). The Air Force Research Laboratory (AFRL), 711 HPW/RHAS Airman Systems Directorate at Wright Patterson Air Force Base, Ohio has demonstrated the efficacy of knowledge management and tracking at the scale of distributed mission training. In 2012-2014, AFRL conducted internal pilot testing on an enterprise KMS for the F-16 tactical fast jet domain. The KMS under test was part of a Category 1 Advanced Technology Demonstration (ATD)

project with Air Combat Command to demonstrate a system that could manage training content for a team of four F-16 pilots (Tannenhill, Portrey, & Watz, 2014).

The environment in which this evaluation took place was a secure testbed area, containing four F-16 Block 40 simulators connected via a Distributed Interactive Simulation (DIS) network. DIS is a widely used data standard for ensuring interoperability between simulation and simulator systems, and the Simulation Interoperability Standards Organization actively maintains the DIS standard (Institute of Electrical and Electronics Engineers, 2012). An Instructor Operator Station (IOS) was used for exercise control and management of the simulators and simulation systems within the environment. This IOS provided a single point of exercise control for the Mission Director. Additional components in this environment included a Computer Generated Forces (CGF) application which produced computer-based adversarial forces that participated in the pilots' training. CGF forces are necessary in the simulation environment to provide opposing forces (OPFOR) against which our warfighters can practice their tactics, techniques, and procedures (TTPs).

The learners in this exercise were retired F-16 operators and all were recruited from an available pool of in-house Subject Matter Experts (SMEs). The average experience of individual SME operators in this exercise was approximately 20 years of operational US Air Force flight time, with most operators having had over one thousand hours of flying time. This level of SME experience provided critical insights into the operational validity of the KMS under test.

The testbed environment contained a library of 270+ scenarios. We define a scenario as the set of instructions, or script, to be executed by the CGF application; essentially, the scenario is what drives OPFOR behaviors. Prior to the KMS evaluation, SME evaluators assessed the level of complexity of all scenarios in the library. A scenario's complexity score indicates the level of cognitive difficulty that scenario presents the 4-ship of F-16's; more entities do not always equate to increased levels of problem solving for the learners. For instance, a large number of OPFOR which do not perform any defensive maneuvering and present themselves as easy targets does not present the same cognitive challenge to a group of learners as does a smaller OPFOR which is aware of blue force tactics and applies counter maneuvering designed to minimize blue force advantages wherever possible.

The KMS test sessions were conducted on several occasions from 2012 – 2014 throughout the KMS's development cycle. These sessions consisted of a Mission Director choosing an initial scenario at a medium complexity level, then allowing the KMS to recommend scenarios based on objective operator performance and subjective inputs from the Mission Director. These limited run testing sessions produced very positive initial results, with the system scaling up complexity level of subsequent scenarios in cases of good pilot performance and ramping down the complexity level in cases of sub-optimal pilot performance. When pilot performance on a particular scenario was in the mid-range, meaning that although mission success was achieved there were mistakes in execution but none of the F-16s were killed during the mission, the system recommended subsequent scenarios of similar difficulty that focused on skills related to the mistakes seen in the mission. Mission Directors observing these training runs generally agreed with the recommendations being generated by the KMS.

KNOWLEDGE MANAGEMENT SYSTEM TAXONOMY

KMS frameworks can generally be summarized into one of three levels, or categories, with each category possessing distinct advantages and disadvantages (Bennett, Berlingo, Tripp, & Watz, 2016). These macro-level categories represent the functionality of knowledge management systems at a macro level; not all KMS systems fall neatly into exactly one of the categories below.

**Table 1. KMS Categories (Bennett, Berlingo, Tripp, & Watz, 2016)**

KMS Category	Advantages	Disadvantages
1. Training Activity Catalogue	- Accessible student status and some content - Calendar tracking - Standard time-based exposure and Assessment - Very efficient	- Fixed content and assessment - One-size-fits-all - Rigid calendar schedule
2. Proficiency By Recurring Activity Tracker	- Adds proficiency/mastery tracking - Potential for "testing out" - Mastery and non-mastery tracked - Typically, content and delivery are static - All "in-system" activity tracked	- Assessments cued to recurring training event schedule - Static content - Non-mastery means completing entire course again - One-size-fits-all
3. Content and Learner Tailoring and Management	- Dynamic content and assessments - Predictive models with exposure and decay functions - Capacity to track learner and content interactions directly - Content delivery based on mastery and gaps - Learner proficiency status reporting	- Complex - Dynamic content and assessments requires maintenance and oversight - May not have key components at needed levels of maturity

Systems traditionally thought of as KMS fall into the first category. These systems and solutions, also known as Training Activity Catalogues, store individual training records data housed in flat files or in databases. Data stored in these systems most often is temporally-based data, such as quarterly or annual training records. As an example, contractors within the Department of Defense complete annual training on multiple subjects. These systems track completion of training and notify users when the next annual training is due. Content delivery within this category is not tailored to individuals and does not adapt to learner styles.

The second category of KMS systems extends functionality from the first and includes some flexibility in how content is delivered to the learner. Examples of this system include KMSs that deliver periodic refresher trainings with an option to skip the learning material and go straight to an assessment. These systems are more dynamic than the Training Activity Catalog; however, all instructional content and tracking is done within a single system.

The third category of KMS systems are composite system-of-systems designs. KMS instances in this category typically integrate a number of existing components operating at high Technology Readiness Levels (TRL). These systems are complex in nature, and it often requires years to develop the individual components and capabilities within the KMS ecosystem. It is advantageous to develop the individual components in parallel with each other to take advantage of common development libraries, code bases, and project knowledge in order to maximize interoperability and efficiency within the system.

These systems take a holistic approach towards knowledge management and performance tracking and include functional areas such as performance measurement and assessments, data warehousing, data recording and playback, linkage of measurements and assessments with instructor and learner feedback mechanisms, and training content selection and delivery. Additionally, to support the need for a KMS to manage, deliver, and administer training content, these systems may also include exercise management capabilities such as interacting with simulators and simulations in a distributed training environment. A pre-requisite to the development of a KMS ecosystem such as the one shown below is the establishment of a scientific vision for how the system will manage learning and training across a set of specified environments.
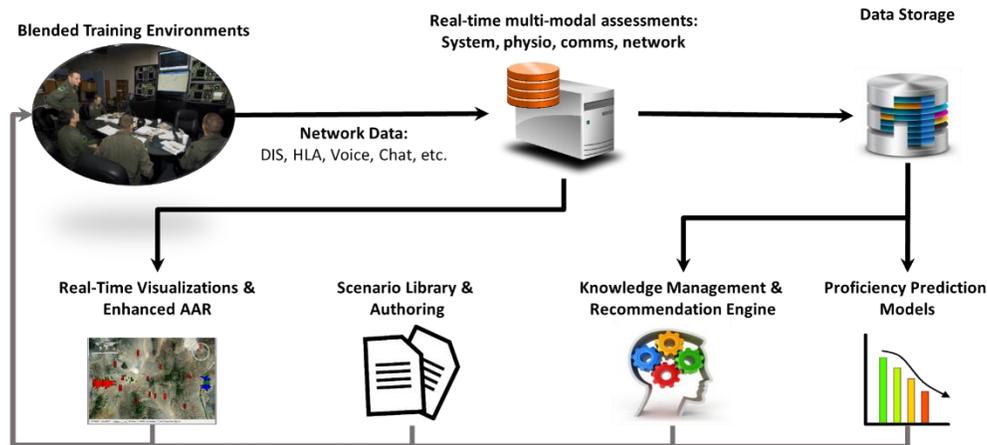
Figure 1. Exemplar Managed Learning Ecosystem

SYSTEMS ARCHITECTURE AND COMMON DATA FORMATS

The complexity and diversity of mission requirements across the spectrum of military training operations requires a system architecture that is extensible, open, and standards-based. It requires scientific processes and methodologies that are generalized in nature, and at the same time easily tailored to the specifics of an environment. It requires a systems architecture that is not wholly dependent upon any particular component and one in which components may be updated without adversely affecting other components. An asynchronous message-based software system architecture supports construction of a modular system using well-defined systems interfaces that specify how components will communicate with other components (Hohpe & Woolf, 2004). System designers separate services within the ecosystem based on a particular function's roles and responsibilities, which can reduce effort when designing the linked interfaces for that function.

Given the quantity of multi-modal, multi-format training data present in a category 3 Knowledge Management System, a common reusable data format and warehouse are pre-requisites. Approaches to this "data stove piping" challenge, such as those referenced in (Neubauer, Freiman, Schroeder, Starrett, Schreiber, & Watz, 2017), (Hruska, Amburn, Long, Kilcullen, & Poeppelman, 2014), and (Watz, DeFabritis, & Hargrove, 2010), must define a common data model, provide for translation from measurement tool specific data formats into the common data model, and support efficient queries to locate the required data. Specifications from the e-learning world, for example the Experience API, are working to provide standardized methods and common data models which seek to enable interoperability between various learning management systems.

PROFICIENCY PREDICTION AND TRAINING MANAGEMENT

Collecting performance data in a common, central format is a necessary, but relatively simple, task compared to deriving actionable human proficiency predictions from the data warehouse. This section outlines the requirements and processes for predicting proficiency and applying those predictions to improve training efficiency and effectiveness. A successful proficiency prediction system must adapt to the available data, fuse data from multiple sources, and be data-driven (Neubauer, Freiman, Schroeder, Starrett, Schreiber, & Watz, 2017).

Data collection is likely inconsistent within real-world training systems, and a successful proficiency prediction system must adapt. In any one training event, the trainees may not employ all of the available measurement tools, perhaps due to budget, schedule, and/or technical constraints. The available measurement tools will therefore evolve with training needs and available budgets. As contracts end, personnel transition to other projects, and priorities shift, some measurement tools ultimately fall out of use and new tools enter the training environment. The proficiency prediction approach must not depend on specific measurement tools and must automatically adapt to the available data sources.

No one measurement tool provides a complete picture of proficiency and a successful proficiency prediction system must fuse multiple objective and subjective data sources spanning multiple modalities. For example, an objective, system-based tool might measure "range to threat" and calculate risk. Assigning a "good" or "bad" assessment to that condition requires context and expert judgement, system-based data, and multiple assessments combined to predict proficiency accurately. Additionally, multiple data sources provide concurrent validity that adds confidence about the accuracy of results. In general, a single measurement tool or modality measures proficiency on portions of selected performance constructs, and it is therefore essential to utilize multiple, complementary measurement tools. The proficiency prediction approach appropriately draws on multiple data sources to create a comprehensive overview.

Some approaches depend only on subjective expert judgement and criteria, and objective data must drive a successful proficiency prediction system. Expert judgements can be subject to variability, and they often depend on training system capabilities and evolving instructional standards. For example, a radar upgrade in a flight simulator may contribute towards observed changes in the performance of the pilots. Rather than requiring a refresh of expert-provided judgements, a data-driven approach applies machine learning techniques to the data to determine proficiency standards, decay rates, training effects, and predictions. The system is then able to function with missing or incomplete data and report increasing confidence as more data becomes available. The proficiency prediction approach must therefore be capable of assimilating and integrating expert judgements as one of many data sources.

Performance Constructs

Many human performance measurements, especially those that are objective, are not comprehensive enough to represent proficiency usefully in isolation, and a mapping from measurements to performance constructs is necessary. Instructors and students often understand proficiency based on higher-order performance constructs. For example, a "range to target" measurement is too low-level, and a higher-order concept like "weapons employment," which combines many measurements, contexts, and assessments from multiple data sources provides a more useful perspective on proficiency (Stacy, Merket, Puglisi, & Haimson, 2006). Higher-order performance constructs are often derived from training standards. Within the USAF fast jet community, the Mission Essential Competency (MEC^{tm}) framework provides these performance constructs, but other existing training standards may provide the constructs as well (Colegrove & Alliger, 2002).

Typically, the process for developing measurements begins with training standards or related performance constructs, so training system designers have some data about the relationships between measurements and performance constructs. In cases where the mapping data is unknown or incomplete, surveys of experts are the next best option. The mapping process involves assigning an assessment or score to each measurement and aggregating multiple measurements into an estimate of proficiency on each performance construct.

Assignment of assessments or scores to measurements depends on the measurement and context (Stacy, Merket, Puglisi, & Haimson, 2006). For example, a pilot might accept increased risk given the expected consequences of mission failure. A criterion-based approach applies expert-derived rules and cut scores to translate a measurement into an assessment. Given enough historical data and a well-defined concept of "mission success", the criterion-based approach could build a model that predicts probability of mission success, where this probability serves as a proficiency score, with measurements as predictors. A norm-referenced approach compares individual performance to peers in similar missions. Norm-referencing allows the system to distinguish "average", "better than average", and "worse than average" performance without expert judgement. The norm-referenced approach works well when criteria are complex and evolve over time (Neubauer, Schroeder, & Freiman, 2017).

A weighted mean may aggregate multiple measurements, both subjective and objective, into per-construct proficiency scores. While simple, (Neubauer, Schroeder, & Freiman, 2017) found the approach effective. Initially, expert judgements provide the weights, and machine learning techniques may automate weight selection as the amount of data available increases. Supervised learning algorithms could incorporate expert judgements and allow the system to improve with use.

Team-based training complicates proficiency prediction since it is difficult to separate individual proficiency from effects of other team members. For example, a single low-performing team member could affect the performance of the team, or a high-performing team member could mask the effects of low-performing team members. Thus, proficiency prediction must adjust for the "teamness" construct when predicting individual proficiency (Arthur, 2012).

Training environments consist of formal training with performance measurements, formal training with limited performance measurements, and informal training events during latency intervals that moderate skill decay. A proficiency prediction system must account for the latency inherent in interval training. Ideally, the system would track such training, but the informal nature of the training often makes tracking difficult. In the absence of latency training records, the system must be capable of predicting the impact of typical latency training.

Proficiency Modeling

Proficiency prediction requires two models, one that predicts the decay of proficiency on the performance constructs and a second that predicts the impact of training on proficiency. Together, these models allow prediction of past proficiency where there was insufficient human performance measurement and future proficiency given a hypothetical training schedule, as depicted in Figure 2. The decay model predictors include predicted proficiency on each performance construct prior at the conclusion of the last the training event, trainee demographics, trainee experience, and latency training events. The impact model predictors include predicted proficiency on each performance construct prior to the training event, trainee demographics, and trainee experience (Neubauer, Schroeder, & Freiman, 2017).
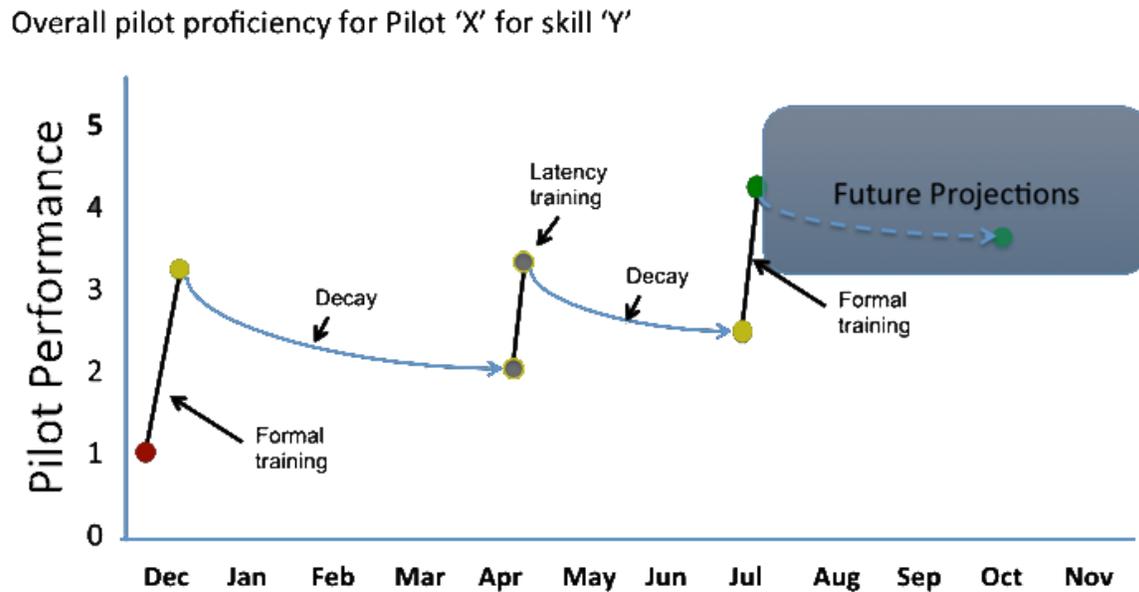

Figure 2 - Conceptual Training Sequence

Training Management

Training management builds on the proficiency modeling with decision support capabilities for students, instructors, and leadership (Neubauer, Freiman, Schroeder, et al., 2017). This capability is the end goal that enables the "content delivery based on mastery and gaps", individualized learning, and adaptive training, key features of a category 3 KMS. Training management assesses current, and predicts future, proficiency of teams and individuals based on curriculum and readiness standards established by USAF and Coalition partner guidelines for the optimal live flight to virtual mix ratio for experience and inexperienced warfighters (Colegrove, Rowe, Alliger, Garrity, & Bennett, 2009). Thus, the KMS recommends the optimal training schedule given training priorities, proficiency goals, and resource scheduling. Training management may function on multiple timeframes. Short-term or inner-loop recommendations, often referred to as intelligent tutoring or intelligent adaptation, provide feedback to trainees during the training event on the seconds to hours timeframe. Middle-loop recommendations provide feedback between training events, typically on the hours to weeks timeframe. Outer-loop recommendations support long-term organizational and career goals, typically on the weeks to years timeframe (Serfaty, 2017).

Baldwin and Ford (1988) define transfer of training as, "the degree to which trainees effectively apply the knowledge, skills, and attitudes gained in a training context to the job" (para. 1). Thus, the ultimate goal of any instructional program is to positively influence on-the-job performance and promote skill maintenance. Unfortunately, research has demonstrated that a substantial amount of the return on investment (RoI) for employers is lost as what is trained either

fails to translate to the work environment or decays sharply with the passage of time (Lim & Morris, 2006; Saks & Belcourt, 2006; Burke & Hutchins, 2007). Despite the elusiveness of tangible transfer, the field of training research has provided several key findings for methods and techniques that can be used to maximize the degree of transfer in organizations, which is especially relevant to those that are employing a KMS (Ford & Weissbein, 1997; Alliger, Tannenbaum, Bennett, & Traver, 1998; Cheng & Ho, 2001; Arthur, Bennett, Edens, & Bell, 2003; Burke & Hutchins, 2007; Blume, Ford, Baldwin, & Huang, 2010; Baldwin, Ford, & Blume, 2017).

Prior to training, it is critical to identify and fully understand the requisite task requirements and competencies that are essential to job performance as they align with organizational and operational goals (Cheng & Ho, 2001; Burke & Hutchins, 2007). The AFRL KMS system described above fulfills this requirement through the utilization of the MEC[tm] framework to establish the core competencies associated with all relevant personnel positions (Colegrove & Alliger, 2002). As training is being conducted, it is recommended that simulations with high psychological fidelity be employed to enable ample practice opportunities with appropriate feedback regarding successful and unsuccessful performance (Salas, Kraiger, & Smith-Jentsch, 2012). Also, the fusion of performance and system data to make informed decisions about actual trainee learning and skill development (via KMS) is a key factor (Baldwin, & Huang, 2010; Baldwin, Ford, & Blume, 2017; Bennett, Tripp, Berlingo, & Watz, 2016). The AFRL KMS satisfies these recommendations through continuous fidelity evaluations on all platform and operator simulation environments and the meta-data processing, archival, and reporting systems described above (Schreiber & Bennett, 2006).

Guidance is also provided on methods to foster transfer and support learners as they transition into their on-the-job environments once training is completed (Lim & Morris, 2006). Common recommendations included approaches such as utilizing data-driven debriefs and after-action reviews (AAR) of training performance (Salas, Kraiger, & Smith-Jentsch, 2012). Leveraging the capabilities provided by AFRL's Live Virtual and Constructive Network Control Suite (LNCS) toolset, the AFRL KMS system provides learners with a robust debrief and AAR capability including in-depth performance data tracking, and replay options to foster trainee skill development and review of valuable lessons-learned. Research has also recommended that organizations implement periodic evaluation and content refresh cycles examining training relevance and realism as business and operational needs evolved over time (Axtell, Maitlis, & Yearta, 1996; Grossman & Salas, 2011; Baldwin, Ford, & Blume, 2017). The AFRL KMS includes MEC[tm] core competencies that are periodically updated and refreshed to meet the fast-paced and dynamic operational landscapes of modern military enterprises (Colegrove & Alliger, 2002; Schreiber & Bennett, 2006), which allows the system to provide relevant, realistic, and empirically-based skill and knowledge development experiences in an optimal and efficient manner to warfighters.

CONCLUSIONS AND LESSONS LEARNED

Real world training environments are complex and variable, which is the source of most challenges developing and deploying a KMS capable of providing adaptive training for individuals and teams of teams. All category 3 KMSs face similar challenges integrating multiple complex systems. First, performance measurement and assessments tools change, and a KMS must not depend on any one data source. Early versions of the AFRL KMS included an internal database designed around one core data source with optional extensions from other data sources, but this assumption failed when deploying the KMS to new training environments. The adoption of a more SOA-like approach will increase flexibility, permit data sources from new domains to be integrated quickly because a common specification exists and support a more general application base. Second, warehousing data in a common format continues to be a challenge due to the quantity of data and difficulty identifying the appropriate level of detail, such as the periods for summary data and time-sampled data frequencies. Third, areas of responsibility for different components within the system, such as control of simulator initialization, is a challenging area for systems engineers. In general, a KMS that is an overlay to existing training systems with minimal changes to existing processes is more likely to achieve adoption.

In today's increasingly data rich training and operational environments, the ability to gain insights into trends, trajectories, and needs has increasing potential to inform decision making and provide solutions to what have been seemingly intractable problems. Two essential challenges are to first integrate disparate sources of data, and to second create a system that can, in an automated fashion, combine those data sources into meaningful representations of current and future human performance. The path to address these challenges in the future will be the same as it is now: to create a common format and to automate the processes that derive meaning from the data, given the quantities and data that will be included. Challenges that will become more salient as work in this field progresses fit into four

categories: technological, legal, philosophical, and cultural. Technological issues include building infrastructure for data sharing and interoperability at a broader system-wide level, which will likely lead to legal challenges regarding data rights and ownership. There is also the integration challenge of ensuring that any new capability is compatible, at some level, with existing enterprise tools and processes to be successful. The philosophical challenge will be to make sure that the indicators and data points that we select truly represent those things that we are after and help lead us to the outcomes we want. The cultural challenge will be for trainers to incorporate KMS data and trust the recommendations enough to make them a part of their training process.

REFERENCES

Alliger, G.M., Tannenbaum, S.I., Bennett, W., Jr., & Traver, H. (1998). A meta-analysis of the relations among training criteria (Report # AFRL-HE-BR-TR-1998-0130). Brooks Air Force Base, TX: United States Air Force.

Arthur, W., Jr., Bennett, W., Jr., Edens, P.S., & Bell, S.T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. Journal of Applied Psychology, 88, 234-245.

Arthur, W., Jr., Glaze, R.M., Bhupatkar, A., Villado, A.J., Bennett, W., Jr., and Rowe, L.J. (2012). Team task
analysis: Differentiating between tasks using team relatedness and team workflow as metrics of team task interdependence. Human Factors, 54(2), 277-295.

Axtell, C.M., Maitlis, S., & Yearta, S.K. (1997). Predicting immediate and longer-term transfer of training. Personnel Review, 26(3), 201-213.

Baldwin, T.T., & Ford, J.K. (1988). Transfer of training: A review and directions for future research. Personnel Psychology, 41, 63-105.

Baldwin, T.T., Ford, J. K., & Blume, B.D. (2017). The state of transfer of training research: Moving toward more consumer-centric inquiry. Human Resource Development Quarterly, 28(1), 17-28.

Bennett, W., Berlingo, M., Tripp, L., & Watz, E. (2016). Creating Systems to Actually Manage Learning: Creation of a Service Oriented Architecture. Proceedings of the Simulation Interoperability Workshop – Fall 2016, 16F-SIW-043.

Blume, B.D., Ford, J.K., Baldwin, T.T., & Huang, J.L. (2010). Transfer of training: A meta-analytic review. Journal of Management, 36(4), 1065-1105.

Burke, L.A., & Hutchins, H.M. (2007). Training transfer: An integrative literature review. Human Resource Development Review, 6(3), 263-296.

Cheng, E., & Ho, D. (2001). A review of transfer of training studies in the past decade. Personnel Review, 30(1), 102-118.

Colegreove, C. & Alliger, G. (2002). Mission Essential Competencies: Defining Combat Mission in a Novel Way. NATO RTO Studies, Analysis and Simulation (SAS) Panel Symposium. Brussels, Belgium.

Colegrove, C., Rowe, L.J., Alliger, G., Garrity, M., & Bennett, W., Jr. (2009). Defining the Training Mix – Sorties, Sims, and Distributed Mission Operations. Paper No. 9166. Proceedings of the 2009 Interservice/Industry Training, Simulation, and Education Conference. Orlando, FL.

Davenport, T.H., & Prusak, L. (2000). Working knowledge: How organizations manage what they know. Cambridge, MA: Harvard Business School Press.

Falvo, D.A., & Johnson, B.F. (2007). The use of learning management systems in the United States. TechTrends, 51(2), 40-45.

Ford, J.K., & Weissbein, D.A. (1997). Transfer of training: An updated review and analysis. Performance Improvement Quarterly, 10(2), 22 – 41.

Goodwin, G., Johnston, J., Sottilare, R., Brawner, K., Sinatra, A., & Graesser, A. (2015). Individual learner and team modeling for adaptive training and education in support of the U.S. Army learning model: Research outline (Report # ARL-TR-0336). Aberdeen Proving Ground, MD: United States Army Research Laboratory.

Grossman, R., & Salas, E. (2011). The transfer of training: What really matters. International Journal of Training and Development, 15, 103 – 120.

Hohpe, G., & Woolf, B., (2004). Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions. Boston, MA: Pearson Education.

Hruska, M., Amburn, C., Long, R., Kilcullen, T., & Poeppelman, T. (2014). Experience API and Team Evaluation: Envolving Interoperable Performance Assessment. Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC). (Paper #14157). Orlando, FL.

Institute of Electrical and Electronics Engineers. (2012). IEEE Standard for Distributed Interactive Simulation - Application Protocols. IEEE Std 1278.1™-2012. New York, NY.

Lim, D.H., & Moris, M.L. (2006). Influence of trainee characteristics, instructional satisfaction, and organizational climate on perceived learning and transfer of training. Human Resource Development Quaterly, 17(1), 85-115.

Neubauer, P., Freiman, M., Schroeder, M., Starrett, J., Schreiber, B., & Watz, E. (2017). Predicting Pilot Proficiency with Interoperable Human Performance Assessment. Proceedings of the Simulation Interoperability Workshop – Winter 2018, 17-SIW-003.

Neubauer, P., Schroeder, M., & Freiman, M. (2017). Predicting, Analyzing, and Tracking Training Readiness and Needs – Final Report. Woburn: Aptima, Inc.

Saks, A.M., & Belcourt, M. (2006). An investigation of training activities and transfer of training in organizations. Human Resource Management, 45, 629-648.

Salas, E., Tannenbaum, S.I., Kraiger, K., & Smith-Jentsch, K.A. (2012). The science of training and development in organizations: What matters in practice. Psychological Science in the Public Interest, 13(2), 74-101.

Schreiber, B.T., & Bennett, W., Jr., (2006). Distributed mission operations within-simulator training effectiveness baseline study: Summary report. (Report # AFRL-HE-AZ-TR-2006-0015-Vol I). Mesa, AZ: Warfighter Readiness Research Division, United States Air Force.

Serfaty, D., (2017). Learning at Scale: Can Data Analytics Help Us Push the Boundaries of Human Performance? Maritime Interest Training Support Group, Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC). Orlando, FL.

Stacy, E., Merket, D., Puglisi, M., & Haimson, C. (2006). Representing context in simulator-based human performance measurement. Paper No. 2856. Proceedings of the 2006 Interservice/Industry Training, Simulation, and Education Conference. Arlington, VA: National Training and Simulation Association.

Tannenhill, B., Portrey, A., & Watz, E. (2014). Developing a Learning Management System for LVC and DMO Environments. Proceeding of the Simulation Interoperability Workshop – Fall 2014, 14F-SIW-043. Orlando, FL.

Watson, W.R., & Watson, S.L. (2007). What are learning management systems, what are they not, and what should they become? TechTrends, 51(2), 40-45.

Watz, E., DeFabritis, D., & Hargrove, J. (2010). Unlocking Data: Realizing the Hidden Potential of Discrete Data Formats. Proceeding of the Simulation Interoperability Workshop – Europe 2010, 10E-SIW-032. Orlando, FL.