Precision Training in Surface Combatant Systems: Teaching the Student, Not Just the Class

Janet Spruill
Aptima, Inc.
Orlando, Florida
JSpruill@aptima.com

Jeffrey M. Beaubien, Ph.D.
Aptima, Inc.
Woburn, Massachusetts
JBeaubien@aptima.com

Lloyd Kleinman
USN Center for Surface Combat Systems
Dahlgren, Virginia
Lloyd.Kleinman@navy.mil

ABSTRACT

The U.S. Navy continues to make advances in data-rich live, virtual, and constructive (LVC) training environments that provide the ability to harness human performance data and apply machine learning (ML) and artificial intelligence (AI) to realize gains in mission performance and readiness. Additional advances in instructional methods and data science, when combined with this ability to understand and act on human performance data, have created a foundation for creating precision learning environments.

During live training and tutoring events, instructors offer hints and guidance in response to the students' verbal and nonverbal cues. They may also modify the sequence of training content and direct the student to additional practice time or skill remediation activities, as required. Precision learning technologies aim to emulate, not replace, this guidance in order to provide an optimal, tailored learning experience for every student. These technologies rely on real-time measures of learner performance and use algorithms that determine precisely what the learner knows in order to recommend what learning experiences should occur next. By tailoring the sequence, difficulty, and type of learning content to the needs of each individual student, precision learning approaches can accelerate time to proficiency. Further, when critical data and performance indicators are captured and catalogued, they can be used in individual and team assessments across domains, in after action reviews, and as a means of tracking performance and proficiency over time.

This paper details the precision training concepts and technologies recently adopted by the U.S. Navy's Center for Surface Combat Systems (CSCS) International Programs (IP) Office. The authors describe methods used to create the first implementation of an environment that enhances existing training content, delivers an optimized learning path, and helps instructors know exactly how each student is performing. Specific guidelines and lessons learned are shared so that readers can implement these approaches in their own organizations.

ABOUT THE AUTHORS

Janet Spruill is Vice President of Programs at Aptima, Inc. She has more than 30 years of experience leading modeling, simulation and training (MS&T), and organizational development programs across defense, federal, and commercial sectors. She has deep expertise as a human performance technologist, learning solutions architect, and principal consultant developing and working with technology-based learning systems. As VP, Janet guides business strategy planning and execution, supports successful transition of science and technology (S&T) research programs, and leads Aptima's Orlando operations.

Lloyd Kleinman is the Chief Technologist for the Security Assistance/International Programs Directorate of the U.S. Navy's Center for Surface Combat Systems located in Dahlgren, Virginia. He has served in the U.S. Navy and has worked for over 30 years with the Aegis Combat System. He is responsible for the research and development of training tools that include PC-based simulation, distributed learning systems, as well as embedded training programs for the Aegis Combat System.

Dr. Jeffrey M. Beaubien is a Distinguished Principal Scientist in Aptima's Learning and Training Systems division. For the past 20 years, his work has focused on training and assessing leadership, teamwork, and decision-making skills. His research has been sponsored by the U.S. Navy, the U.S. Army, the U.S. Air Force, and the Telemedicine and Advanced Technologies Research Center (TATRC), among others. Dr. Beaubien holds a Ph.D. in Industrial and Organizational Psychology from George Mason University, a M.A. in Industrial and Organizational Psychology from the University of New Haven, and a B.A. in Psychology from the University of Rhode Island.

Precision Training in Surface Combatant Systems: Teaching the Student, Not Just the Class

Janet Spruill
Aptima, Inc.
Orlando, Florida
JSpruill@aptima.com

Jeffrey M. Beaubien, Ph.D.
Aptima, Inc.
Woburn, Massachusetts
JBeaubien@aptima.com

Lloyd Kleinman
USN Center for Surface Combat Systems
Dahlgren, Virginia
Lloyd.Kleinman@navy.mil

INTRODUCTION

For the past several years, the U.S. Navy (USN) has been leveraging rapid acquisition authorities provided by Congress in the fiscal year 2016 National Defense Authorization Act. These provisions, and the subsequent OPNAV detailed instructions, were aimed at helping to fast track the development of new weapon systems and to speed up the acquisition cycle so new capabilities could be fielded sooner — years sooner in many cases. But as acquisition cycles have shortened for new missile, laser, and unmanned systems, having the qualified personnel available who can expertly maintain and employ them has become even more challenging. According to Mr. James Geurts, Assistant Secretary of the Navy for Research, Development and Acquisition, "whether it's … building ships faster or putting new capabilities on ships, eventually we will be limited in its effectiveness by how fast we can train the crew and make them proficient" (Eckstein, 2019).

To address this critical need, the USN is investing heavily in advancing training methods and technologies focused on accelerating Sailor time to proficiency to help drive operational readiness. Schoolhouse training efforts are shifting away from a focus on training *completion* to one focused on performance-based *outcomes*. As a result, the need to measure and verify Sailor proficiency has never been more important. The Ready, Relevant Learning (RRL) Pillar of the Sailor 2025 (S2025) initiative calls for the delivery of learning content at the point of need, as well as for the tracking and analysis of the *effectiveness* of training and instruction. USN initiatives such as S2025 RRL and the Surface Training Advanced Virtual Environment (STAVE) are seeking to increase training effectiveness through blending traditional classroom learning with highly detailed 3D virtual tools and virtual and instructor-led labs where students can train, practice, and gain proficiency in realistic and progressively more complex scenarios.

Even with this shift to a performance-based model, most training still focuses on "teaching to the class"; progressing an entire group of students through the schedule using a one-size-fits-all approach that that cannot fully consider each student's individual learning experiences or outcomes, nor the wide variation in student knowledge and skill levels. There is a growing interest in USN leadership to explore how artificial intelligence (AI) can help speed up training and validate proficiency as new capabilities are rapidly fielded. To date, however, most AI-based training efforts across the U.S. military have been limited to research-based projects or prototype demonstrations. By comparison, the International Programs (IP) Directorate at the USN's Center for Surface Combat Systems (CSCS) has embraced AI-based training for a growing portion of its foreign military sales (FMS) Aegis curriculum to precisely target training content, accelerate learning, and validate student proficiency as new capabilities are fielded or equipment baselines are upgraded. CSCS IP began implementing AI-based "precision" training in mid-2018, and in early 2020 expanded that focus into a broader *Precision Training Strategy*.

PRECISION TRAINING OVERVIEW

Precision training has emerged as the convergence of 1) personalizing and tailoring training at scale, 2) adapting and optimizing a student's learning path, 3) emulating rich human feedback, and 4) transforming the student from a passive recipient of information to an active participant in their learning process. It leverages AI and other precision technologies to allow training to be *exact* and *accurate*, using a "surgical focus to deliver exactly the right content in the right way in order to create learning *flow* for each student" (Serfaty, 2019).

During one-on-one human tutoring, instructors offer hints and guidance in response to the students' verbal and nonverbal cues. They also modify the sequence of training content and direct the student to additional practice time or skill remediation activities, as required. In distributed and virtual learning environments, advanced learners can waste precious time reviewing material that they have already mastered, whereas marginal learners do not receive the additional remediation that they need to meet the course proficiency standards. Precision training seeks to address this need.

It has long been recognized that tailoring the presentation of training to the needs of individual students leads to massive gains in student outcomes (Bloom, 1984). However, adapting the presentation and sequencing of training content for an individual student has traditionally been the domain of instructors, making it logistically challenging to do at scale. Precision training technologies can do this at scale through real-time, unobtrusive measures of student performance and by using algorithms that determine precisely what each student knows in order to recommend what tailored learning experiences should occur next.

Using human performance measurements, combined with machine learning (ML) and AI, precision training finds that "sweet spot" where instruction is most beneficial for each student — just beyond his or her current level of capability. This is known as the "Zone of Proximal Development" (Vygotsky 1978). To train to the student's sweet spot requires that we first define and quantify acceptable levels of performance, compare measured student proficiency to that benchmark, and then have a learning ecosystem in place to precisely target those areas that are in greatest need of remediation. In naval maintenance training, for example, students must master a broad range of fault isolation and corrective maintenance tasks. Because faults vary in type and difficulty, students can learn faster by having the system tailor training to the types of faults that they have not yet mastered, and by selecting faults at the appropriate level of difficulty.

Precision training is data driven and science anchored. It is deeply rooted in the science of learning and is enabled through human performance measurement and providing high-quality feedback to students during blended learning activities. Importantly, it aims to emulate, not replace, live instructor/mentor guidance. Although conventional learning theory states that it takes 10,000 hours (or roughly 10 years) to achieve domain expertise, providing tailored, deliberate practice with relevant feedback can greatly accelerate the time it takes novices to perform at a high level of proficiency (Ericsson & Charness, 1994; Ericsson, Krampe, & Tesch-Romer, 1993).

CSCS INITIATES AI-POWERED PRECISION TRAINING

The mission of CSCS IP is to support U.S. security interests by helping to build a well-trained and ready international network of allies and partners who employ Aegis-equipped ships. CSCS is rapidly fielding sophisticated simulation technology to reduce reliance on physical hardware and provide a more flexible approach at lower cost. CSCS IP leadership was interested in applying advances in AI, together with advances in learning science, to these training technologies to move beyond the current one-size-fits-all training approach — and to better align with future plans to increase in-country training to our allied partners using distributed learning. In mid-2018, CSCS IP initiated its first precision training project to implement AI-powered training in an effort called *Simulator-Harnessed Intelligent Performance Measurement and Adaptive Training Environment* (SHIPMATE).

SHIPMATE was designed as an asynchronous intelligent lab environment that scaffolds learners as they apply knowledge from prior classroom and live lab training to Aegis maintenance tasks from very basic to advanced troubleshooting. Sailor readiness and proficiency is highly dependent on the frequency and recency of experience in realistic training environments that provide both a meaningful context and complex operational environment representations. So, this intelligent system was designed to enhance, or "wrap around," existing virtual simulations to create a fully adaptive, precision training environment. Early on, it was determined that this same "simulator-agnostic" approach could be applied to any training system (e.g., interactive courseware, distributed training exercise, immersive simulation, part-task simulator), where it is not dependent on a specific training delivery method or type of content.

Because many CSCS training systems already use the Experience Application Program Interface (xAPI) to capture measures of learner performance (Smith, Gordon, Hayden, & Johnson, 2019), this has provided a streamlined way to capture student actions in the simulator, such as the extent to which students were performing maintenance activities according to standard and in the correct sequence. Additional unstructured data were captured to further enhance available xAPI data as well as to capture the system and environmental data needed to inform meaningful performance measures. Subsequently, real-time assessment of the student measures was used to inform ML algorithms. These algorithms precisely determined the current skill level of each student and recommended the next training scenarios a student was best positioned to attempt. Finally, the presentation of learning content was adapted to the student using their continually assessed individual strengths and weaknesses. Scaffolding was accomplished using *micro adaptation* within a scenario and *meta adaptation* across scenarios (VanLehn, 2006). Through this persistent measurement and assessment cycle, students and instructors were able to validate that skill acquisition had occurred.

Following the successful completion of this first project, CSCS IP is expanding the application of AI-based training to address a broader portion of its Aegis curriculum and has begun exploring how to expand precision training beyond international programs to address USN training requirements.

FIVE ELEMENTS OF THE PRECISION TRAINING STRATEGY

As an early adopter of AI-based training, CSCS IP recognized both the potential and the advantages of adopting a broader, more systematic strategy. In early 2020, CSCS IP initiated a *Precision Training Strategy* to help standardize and expand efforts to provide an optimal, tailored learning experience for every student.

The CSCS IP Precision Training Strategy has five (5) key fundamentals:

1) Unlock the value of training data to develop critical insights about each learner.
2) Measure learner performance to drive meaningful assessment.
3) Use smart algorithms to recommend the optimal, personalized learning path.
4) Encourage students to take control of their learning with the aid of intuitive “proficiency maps.”
5) Blend learning theories and strategies that best support a learner-centered training model.

Unlock the Value of Training Data

Acquiring training data is an important means to an end, but not an end in itself. Analyses of rich sets of untapped data can provide actionable insights from the individual Sailor in a single training event to a class of students or an entire curriculum and can even be aggregated to the schoolhouse. For example, *descriptive* analytics can be used to mine historic student performance data from a program of instruction to identify topic areas where students consistently have the most difficulty. *Predictive* analytics can help forecast how skill acquisition will be impacted with a transition from physical technical training equipment (TTE) to virtual equipment labs. And finally, *prescriptive* analytics can help provide instructors and organizations with data-informed recommendations to support goals such as optimizing student throughput and efficiency, managing resources and training pipelines, and maximizing the return on training investments.

Data are also the fuel of an assessment engine, and persistent, data-rich student assessment is an essential element of precision training. At CSCS, the data being generated during training events come from many sources and exist in many forms. Because data are the basic asset that have the potential to provide analytical insights, they must be ingested, converted to meaningful measurements, and analyzed to add intelligence using algorithms and computational models. The core power behind precision training is the ability to consume data from multiple sources, fuse them into one data model, and allow computations to run on top of those data to provide a better understanding of exactly what is going on in the training environment and why.

In today’s ever-expanding landscape of training data, precision training technologies can help tune out most of the noise to focus on data that are most meaningful and translate them into human performance measures. Modern data standards, such as xAPI, make it easier to harness data in a structured and meaningful format. Many organizations are developing systems and employing xAPI, which may then be used as the primary syntax for measuring student behaviors. These are augmented with additional data, as needed, and the aggregate used to continually assess across skill competencies.

Measure Learner Performance to Drive Meaningful Assessment

Reliable and valid human performance measurement is the foundation of precision training. Precision training systems require an objective (system-based) measurement capability to capture and fuse real-time data streams from simulators, interactive courseware, training exercise environments, and other data sources and to convert those data to meaningful measures of student performance. *Behaviorally based* measures are developed across a wide array of training scenarios to capture the behaviors of a student as a reliable indicator of their knowledge or skill state.

In the Aegis training environment, there is an abundance of generated data, along with many decision points, interactions, and parallel activities. All of these can be measured (independently or in some logical sequence) and then assessments can be made regarding learner performance. In this context, *assessments* are the interpretation of the knowledge or skill state of the student based on the performance data provided by the measurements.

Continuous assessment involves the process of comparing one or more *measured* behaviors of an individual or team to those *expected* and then deciding what needs to happen next. For example, during a live training event, instructors *expect* a student to perform a fault isolation process according to the sequence prescribed in the authoritative technical manual while adhering to safety guidelines. When they *observe* a student performing training activities, there may be deviations from the ideal. Those deviations may be indicators of knowledge or skills deficiencies that can be resolved through feedback, practice, remediation, or some other form of training intervention. Observation of the student enables an instructor to compare the observed to “what right looks like” and thus *assess* the student’s performance.

An *automated* performance measurement capability is crucial for processing large amounts of data — including any combination of xAPI data, unstructured “raw” data, or another standard data format such as the High Level Architecture (HLA) used for distributed simulations — so that the data can be converted to meaningful measures and provide continuous performance assessment. This automated capability becomes increasingly important as training scales to a more distributed model. If a training system generates xAPI statements, as many of the CSCS training systems do, this common data structure can be used to share information about the learner’s context and performance. The project team was able to leverage available xAPI data, together with recent theoretical and technical advances in performance measurement technologies that had resulted from earlier efforts at the Naval Air Warfare Center Training Systems Division (NAWCTSD). In one NAWCTSD use case, unobtrusive system-based measurement technologies are used to calculate Navy P-8A Weapons Tactics Trainer crew-based performance measures (Wheeler, Tindall, Killilea, Tolland & Dean, 2017).

The CSCS IP strategy calls for a structured and repeatable *human performance measurement* development process. First, a set of critical job tasks is defined. Next, indicators of high, average, and low performance on these tasks and underlying skills are developed. Finally, measures are developed to quantify task performance, inform assessments, and provide systematic feedback. These steps are accomplished over a series of workshops in which subject matter experts (SMEs) work closely with behavioral scientists to define specific *performance indicators* (PIs) and measures.

PIs provide a framework on which to develop measures based on critical decisions and events. They are observable behaviors that allows an expert (i.e., someone familiar with mission objectives and task requirements) to recognize whether an individual or team is performing well or poorly. As such, PIs serve as the building blocks for performance measures used in assessments (Wiese, Nungesser, Marceau, Puglisi & Frost, 2007). For an Aegis technician, these PIs might include knowing names and acronyms of Aegis system components, troubleshooting and repair procedures, which technical publication provides information describing specific topics, and correct safety precautions that must be taken to avoid personnel injury and equipment damage. To remain objective, it is important to identify observable rather than inferred behaviors that can be collected through the user’s interactions with a training system.

Although some PIs are readily translated into performance measures (such as whether or not the maintenance technician disabled the power before beginning the maintenance activity), others often require more detailed information to define behaviorally anchored performance measures (such as the extent to which the maintenance technician saved time and other scarce resources by intentionally avoiding diagnostic procedures that are both time-consuming and of little diagnostic value). In this case, the team would need to continue to identify the system-observable behaviors (for use as anchors) relevant to each PI and to elicit examples of good, average, and poor behavior associated with each of the learning objectives and sub-goals (or some finer grained level of analysis, if appropriate).

The purpose of this final step in measure development is to validate that all PIs and performance measures are operationally relevant, as thorough as possible to support the learning objectives, and appropriately worded using the experts’ language and terminology. A validation event is held with SMEs to ensure that PIs and measures have the following characteristics:

- Are observable
- Are appropriate to system-based collection
- Use appropriate rating scale anchors of good, average, and poor behaviors and wording
- Are comprehensive, identifying any additional measures needed to fill gaps in the measurement framework

Use Smart Algorithms to Recommend the Optimal, Personalized Learning Path

Precision training cannot be achieved on a large scale using traditional, non-adaptive approaches. It requires a model-based approach that can take performance measures and (1) intelligently guide the selection of learning content within a training event (i.e., inner loop adaptation); across and between scenarios, lessons and exercises (i.e., middle loop adaptation); and (2) recommend courses and career paths to optimize the path to competence (i.e., outer loop adaptation)(VanLehn, 2006).

Machine learning helps continually harvest student data in real time and use it to inform intelligent analytical models. These models must focus on combining and assessing measurement indicators to help inform decisions about a student's skill level. Multiple models form the mathematical framework necessary to provide pattern recognition and the statistical modeling able to then feed one or more computational algorithms and update the probability distribution of the *belief state* of student skills. That belief state is then used to recommend and adjust the sequence or type of learning content a student receives.

The adaptive training algorithms used in current CSCS efforts support an intelligent mathematical modeling approach called the *Partially Observable Markov Decision Process* (POMDP). The POMDP framework is sequential in nature, that is, belief state optimization takes place over a sequence of time steps, and at each time step a single decision is made. Furthermore, at each time step the configuration of the environment contains some uncertainty (we say that the environment is *partially observable*).

The POMDP is a statistical Bayesian approach to decision planning under such uncertainty (Smallwood & Sondik, 1973). A POMDP extends the classic Markov Decision Process (Puterman, 1994), and is used to solve problems in which there are observable variables (e.g., student measurements captured from a training system) and non-observable variables (e.g., current assessed capabilities of the learner). This approach mathematically models the learner by combining multiple sources of observable information and hypotheses about non-observable information to form an optimized plan called the *POMDP policy*. This policy is referenced by a software *recommendation engine* in real time that recommends the "next best" training event and transitions the learner through a sequence of learning content. Here, "best" is defined as targeting the learner's greatest skill weakness.

The POMDP policy continues to inform decisions about each new training scenario based on a student's demonstrated skills in all the previous scenarios. When the results of the training are sent back to the training plan, the training plan considers the new measures and recommends more personalized content to the trainee. Each student, and therefore each sequence of scenarios, is unique. It is impossible to predict the exact learning content that will be presented because it is based on a student's individual performance. This approach also provides recommendations that are granular enough (i.e., to a specific skill) that scenarios will have enough variability to ensure that a student can only show proficiency if they have a solid understanding of the underlying knowledge and skills supporting the scenario.

The policy can be developed by initially ingesting a corpus of learner records, which can then be statistically analyzed using ML. However, using POMDP provided an advantage to early CSCS IP efforts of not requiring a large corpus of existing data or learner records to inform the models. CSCS IP was able to initially use the input of subject matter expert (SME) ratings as a starter, and then over time (statistically) reduce the weighting of SME input and prioritize actual learner data (**Figure 1**). The POMDP model is a learning model. As such, it becomes increasingly more accurate over time as more and more learner data are collected.

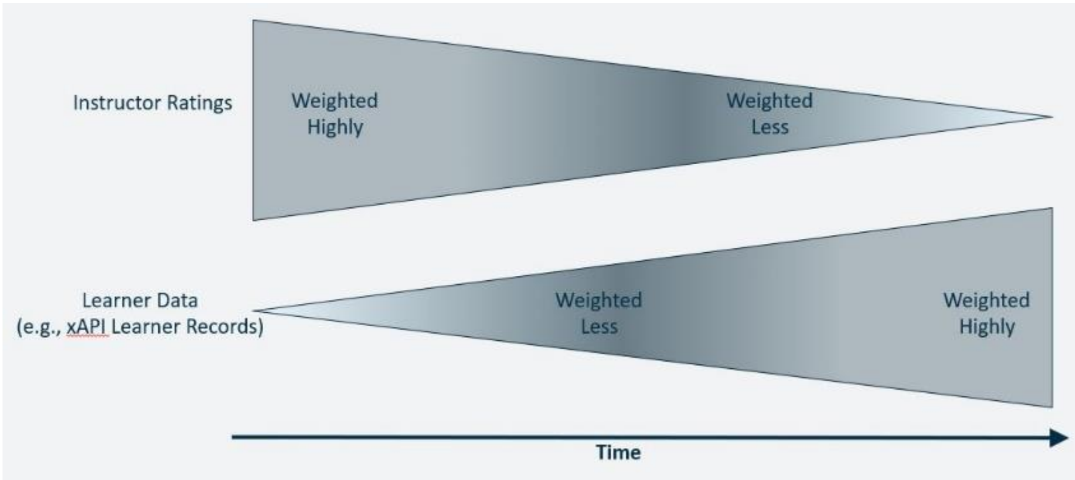

Figure 1. *POMDP relies on more student data over time, which increases accuracy.*

Encourage Students to Take Control of their Learning with the aid of Intuitive "Proficiency Maps"

In order to engage students to become active participants in their training progression, the CSCS IP strategy uses different visualization methods to provide students with a persistent and accurate view of their current knowledge and skill level in the course. Visualizations help depict which subject areas have been mastered and which have not. When a student completes a learning topic, their individual proficiency "map" is updated, showing where they currently are, and where they need to go. This map is represented as a gradient bar chart (**Figure 2**) and helps Sailors immediately draw actionable insights from the large amounts of collected data.

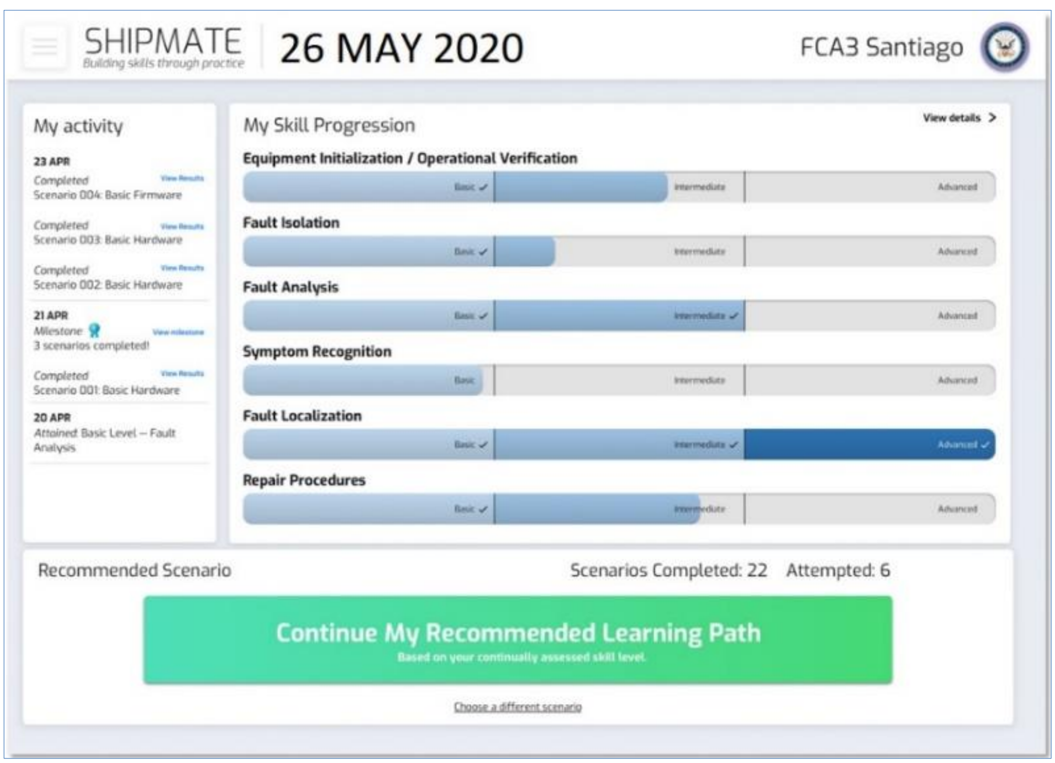

Figure 2. *Proficiency map provides actionable learning insights.*

Even a simple visualization can provide the student with valuable insights within seconds. Each section of the proficiency map allows the student to drill down to more granular detail (**Figure 3**), in order to do the following:

- View course learning objectives mapped to progress
- See alternate ways to filter views of skill progression (e.g., hardware, software, firmware skills)
- View results of completed training activities

At the conclusion of each training experience, students can view a scenario summary screen (**Figure 4**) that provides details about their performance.

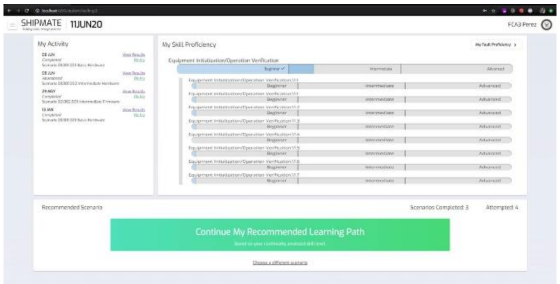

Figure 3. *Students can drill down to view proficiency at the learning objective level.*

Figure 4. *Students can view detailed scoring for each completed training experience.*

Data visualizations like this are critical for presenting raw training data and performance measurements in a personalized and consumable form. This gives students the tools to see directly how their efforts in a portion of training impact their proficiency levels, and it can empower them to take more control of their learning process. It is also consistent with adult learning theories, which view learners as taking an active role in their own self-development (Knowles, 1973).

Visually identifying those skills which are in greatest need of remediation is considered a design best practice. Left to their own devices, many learners continue to practice those skills that are already well developed because success is self-reinforcing (Bandura, 1982). However, in order to provide a well-rounded learning experience, learners must confront those areas that are in greatest need for remediation. Visual summaries depict their unique profile of strengths and weaknesses and can automatically direct them to learning events that will target their greatest areas of weakness (Ericsson, Krampe, & Tesch-Romer, 1993). Finally, this screen provides access to the AI-based *recommender* capability, whereby the system shows the next recommended scenario via the "Continue My Recommended Learning Path" button. The learning scenario associated with this button is continually updated to take the student to content where they have the prerequisite knowledge and skill to successfully complete the troubleshooting scenario.

Blend Appropriate Learning Theories and Strategies

Learning theories provide the foundation for the selection of the best instructional strategies in precision training and allow for reliable prediction of their effectiveness. With AI-based learning, every aspect of how we think about designing and structuring training is up for reimagining. To achieve effective learning outcomes, it is important to link instructional strategies or techniques to the theories of human learning. Several validated learning theories informed the precision training approaches described in this paper.

First and foremost is Knowles' (1973) theory of andragogy, which asserts that — unlike children's education, which is organized into discrete topics and lessons that consist largely of abstract facts and concepts, many of which are not used immediately — adults prefer to link new concepts with their existing knowledge and experiences; focus on real-world, rather than abstract, problems; immediately apply what they have learned; and learn by doing and observing the consequences of their actions. Additionally, the approach was informed by the concepts of distributed practice (Donovan & Radosevich, 1999) and retrieval practice (Roediger & Butler, 2011), which require the learner to actively generate a response, and to do so repeatedly over time, thereby generating the necessary "reps and sets." Finally, recognizing that all reps and sets are not created equal, the personalized learning approach embraces the concept of systematic knowledge engineering by providing the learner with a wide variety of simulated experiences upon which to draw (Klein, Phillips, Rall, & Peluso, 2007).

Because both novices and experts make the same types of logical errors, instructors need to intentionally provide the learners with a broad array of simulated training scenarios — depicting different problems to be solved and different levels of task difficulty) — thereby mimicking the accumulated corpus of experiences that experts draw upon when presented with a new problem (Kahneman & Klein, 2009; Klein, Phillips, Rall, & Peluso, 2007).

LESSONS LEARNED:

The initial CSCS IP precision training efforts described in this paper have resulted in the following lessons learned for maximizing learning-related outcomes:

- **Rich feedback is a critical success factor to increase knowledge and skill retention.** Feedback is an important part of the assessment process. It has a significant effect on student learning and has been described as "the most powerful single moderator that enhances achievement" (Hattie, 1999). The use of real-time data collection and assessments can help target the most appropriate time to provide student feedback, which can impact the transfer of requisite skills (Dideriksen, 2019). Waiting until the end of a training module or course to provide a summative assessment does not leave room to remediate at the time the learning deficiency occurred. Precision training needs to consider making more nuanced formative assessments. This provides just-in-time remediation to students needing additional help, but it also allows more proficient students to progress quickly through the learning content. Such timely diagnostic feedback can only be given when we know the student's weaknesses and strengths at all points in the learning process.
- **Learning complex skills requires considerable practice.** Learning complex military maintenance and operations skills requires more practice and feedback than is typically provided during C School classroom (in-person or distributed) and lab time. Although this is often true in any skills-based training environment, designers must ensure that precision training does not inadvertently choose learning *efficiency* and reduction of learning time over *effectiveness* of learning outcomes. The design of precision training should provide just enough repetitions ("reps and sets") using proximal content (i.e., context and difficulty) to reinforce skill acquisition or remediate any deficiencies. It must also provide enough learning experiences to fully exercise the models and validate skill proficiency, ensuring that skills are mastered before moving forward. An advantage of the precision training environment is that students can get exactly as much practice as they need in order to learn and retain concepts and skills. In most cases, this is much more practice than an instructor could possibly observe or grade.
- **Precision training must also consider and model knowledge decay**. Knowledge decay refers to the loss or decay of trained or acquired knowledge or skills after periods of nonuse (Arthur, 1988). Decay is particularly problematic in situations where individuals receive initial training on knowledge and skills that they may not need to use or apply for extended periods of time. For example, a topic that is mastered during week 2 of a 10-week course may not be retained by week 10 without revisiting and reinforcing at specified intervals. To ensure that topics learned are retained in long-term memory, the precision training system should include techniques and algorithms that periodically reassesses the student. Those results should then be used to adjust the belief state related to that student's knowledge and skill level and then depict that updated level on the student's proficiency map. Because students are constantly revisiting earlier material in new and different ways, a completed proficiency map means that knowledge and skills are less likely to decay and are likely to transfer to the shipboard environment.
- **A larger corpus of learning content increases precision.** A highly personalized learning experience requires content to fully support each student's tailored path. This might include remediation resources such as flash cards, animated process flows, or access to authoritative technical manuals. Plus, if a student is assessed as weak in a skill, the system would need to find a lesson of comparable difficulty to allow additional practice and skill improvement. The adaptive component needs to know the profile of all the content available in order to continually recommend the content that student is ready to learn. And finally, skill validation requires having a sufficient body of content so the student can demonstrate skill proficiency in more than a single scenario

CONCLUSION

The goal of precision training is to meet the exact needs of every student, all the time. By tailoring the sequence, difficulty, and type of learning content to the needs of each individual student, precision learning can accelerate time to initial proficiency, provide rapid upskilling or reskilling in response to a changing environment or threat, and increase efficiency in training pipelines and the use of training resources. Initial efforts at CSCS focused on validating a student's progression toward the acquisition of *technical* skills in a virtual lab environment during schoolhouse training. However, CSCS IP recognizes that when training data and performance indicators are captured and catalogued, they can be used in individual and team assessments across domains and as a means of tracking performance and proficiency over time.

The Precision Training Strategy being implemented at CSCS is directly aligned with and supports the 2018 National Defense Strategy line of effort to rebuild military readiness (U.S. Department of Defense, 1918). It is also aligned with the US Navy's strategic plan, Design for Maintaining Maritime Superiority 2.0 (U.S. Navy, 2018), which mandates a focus on increasing LVC training efforts and "fielding AI/ML algorithms in areas that enhance training."

Looking forward, AI-based learning systems will soon be able to predict the future performance of a Sailor by looking at their training and operational performance over time. It will also be able to study individual learning patterns and preferences, and then help navigate Sailors toward a learning and career path that takes those preferences and strengths into consideration. This kind of intelligence will help the Navy fine-tune training, improve readiness, and make better personnel and manning decisions.

ACKNOWLEDGEMENTS

The authors would like to thank the Center for Surface Combat Systems, International Programs Office, for their continued support of the work described in this paper. However, the contents of this paper are the responsibility of the authors alone.

REFERENCES

Arthur, W., Bennett, W., Stanush, P. & McNelly,T. (1988). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance*, 11(1), 57-101.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*. 37, 122-147.

Bloom, B.S. (1984). The 2 Sigma Problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*.

Dideriksen, A. (2019). AI for military training: Lessons learned from the healthcare industry. *Modern Military Training*. Last accessed June 4, 2020 at: http://modernmilitarytraining.com/training-effectiveness/ai-for-military-training-lessons-learned-healthcare-industry/.

Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84*(5), 795-805.

Eckstein, M. (2019). Artificial intelligence could speed up Navy training as new tech is rapidly fielded. U.S. Naval Institute News (USNI News). Last accessed on May 29, 2020 at: https://news.usni.org/2019/02/21/artificial-intelligence-speed-training-new-tech-rapidly-fielded.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49*(8), 725-747.

Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363-406.

Hattie, J. (1999). *Influences on Student Learning*. Auckland: University of Auckland.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist, 64*(6), 515-526.

Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data–frame theory of sensemaking. In R. R. Hoffman (Ed.), *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making (pp. 113-155)*. New York, NY: Erlbaum.

Knowles, M. (1973). *The adult learner: A neglected species*. Houston, TX: Gulf Publishing Group.

Puterman, M. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming* (1st. ed.). John Wiley & Sons, Inc., USA.

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*(1), 20-27.

Serfaty, D. "Human Performance in the Age of AI," Plenary Address, 2019 Naturalistic Decision Making Conference, San Francisco, June 2019.

Smallwood, R. & Sondik, E. (1973). The Optimal Control of Partially Observable Markov Processes over a Finite Horizon. *Operations Research*. *21*(5), 1071-1088.

Smith, B., Gordon, J., Hayden, T. & Johnson, A. (2020). *Total Learning Architecture: 2019 Report*. Orlando, FL: Advanced Distributed Learning Initiative. Last accessed on May 29, 2020 at: https://www.adlnet.gov/resources/publications/2020/04/2019-Total-Learning-Architecture-Report/.

U.S. Department of Defense. (2018). *Summary of the 2018 National Defense Strategy of the United States of America: Sharpening the American Military's Competitive Edge*. Last accessed on June 1, 2020 at: https://dod.defense.gov/Portals/1/Documents/pubs/2018-National-Defense-Strategy-Summary.pdf.

U.S. Navy. (2018). *A Design for Maintaining Maritime Superiority, Version 2.0*. Last accessed on June 4, 2020 at: https://www.navy.mil/navydata/people/cno/Richardson/Resource/Design_2.0.pdf.

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education*, *16*(3), 227-265.

Vygotsky, L. S. (1978). *Mind in Society: the Development of Higher Psychological Processes.* Cambridge, MA: Harvard University Press.

Wiese, E., Nungesser, R., Marceau, R., Puglisi, M., & Frost, B. (2007). Assessing trainee performance in field and simulation-based training: Development and pilot study results. Paper No. 7160. In *Proceedings of the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)*. Arlington, VA: National Training and Simulation Association.

Wheeler Atkinson, B., Tindall, M., Killilea, J., Tolland, M., & Dean, C. (2017). Standardizing performance measurement for ease of data analytics. In *Proceedings of the 2017 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)* (pp. 2187-2195). Arlington, VA: National Training and Simulation Association.