Learning Analytics with xAPI in a Multinational Military Exercise

Aaron Presnall, Ph.D., Vesna Radivojević
Jefferson Institute
Washington, DC
apresnall@jeffersoninst.org, vradivojevic@jeffersoninst.org

ABSTRACT

As the truism goes, "You can't manage what you don't measure." However, assessing performance in training exercises has classically presented a measurement challenge, made more complex by the paucity of timely, relevant, comparable data on the training audience's performance. Even as the field of learning analytics becomes increasingly sophisticated, military training exercises continue to be assessed in largely subjective and superficial ways. In short, while we may know if the training was completed, it is difficult to objectively answer the basic question: did the exercise do any good?

xAPI is an emerging capability to support learning analytics, but until recently has remained largely untested as a solution for delivering comparable results across complex multi-platform asynchronous learning and performance data feeds at scale. Viking 18, a large multinational civil-military exercise, aspires toward full operational integration of Advanced Distributed Learning (ADL) as an integral part of the exercise experience, including the associated learning analytics supported by xAPI.

This paper presents a case study and lessons learned from the implementation of xAPI in the Viking 18 exercise. It also delivers a summary of the resulting Viking 18 learning analytics, including data from e-learning courses matched against quantitative observation data from the exercise management tool, with the aim of gaining insight on the relationships between training and performance against exercise objectives. As such, we crack open the door to aggregation of exercise performance data in support of operational and strategic planning. Analysis clearly suggests a pattern of enhanced training outcomes by units with higher rates of *Introduction to Viking* (pre-training e-learning) course completion.

ABOUT THE AUTHORS

Aaron Presnall (Ph.D., University of Virginia) is president of the Jefferson Institute and lecturer at the University of Virginia, Department of Politics. He is a political economist specializing in telecommunications regulatory transformation and the relationships of information and participatory decision making. In addition to scholarly works and popular opinion pieces, he has written on the business and political environment of Eurasia for the United Nations Development Program (UNDP), the Economist Intelligence Group, the OSCE and numerous private and governmental organizations in the United States. Before joining the Jefferson Institute, he served with the EastWest Institute for seven years in Prague, then in Belgrade for three years as EastWest's Regional Director of Southeast Europe.

Vesna Radivojevic (MS Mathematics, University of Belgrade) is Southeast Europe Regional Director at the Jefferson Institute. Ms. Radivojevic has over a decade of experience in software architecture and data analytics. She leads a team of programmers and designers, which have developed cutting edge technology projects for the National Science Foundation, PBS Newshour, and Sochi Olympics.

Learning Analytics with xAPI in a Multinational Exercise

Aaron Presnall, Ph.D., Vesna Radivojevic
Jefferson Institute
Washington, DC
apresnall@jeffersoninst.org, vradivojevic@jeffersoninst.org

INTRODUCTION

This paper presents a case study and lessons learned from the implementation of xAPI in Viking 18 – a multinational Computer-Assisted Exercise (CAX). xAPI, or the Experience Application Programming Interface, is a data interoperability specification that facilitates more granular and interoperable human performance data collection. It can enable improved aggregation and analysis of trainee performance data across various technology-supported learning activities.

We also include results from the xAPI-enabled learning analytics used in Viking 18, including data from the e-learning courses matched against quantitative observation data from the exercise management tool, with the aim of gaining insight on the relationships between training activities and outcome performance against exercise objectives. As such, we crack open the door to the aggregation of exercise performance data in support of operational and strategic planning, including training remediation, future exercise design, readiness estimations, and return on investment analyses.

Our effort tested the viability of implementing xAPI across a set of diverse e-learning courses. It also evaluated the utility of xAPI-supported learning analytics in training settings, demonstrating that increased visibility into the training audience's performance yields valuable insights. Finally, analysis of the collected data empirically demonstrated improved training outcomes due to the operational integration of Advanced Distributed Learning (ADL) into the event.

Problem Statement

Strategic-level guidance on training and education in the security sectors has entered a new (r)evolutionary phase. This reflects an increased need to more effectively prepare personnel for a range of complex and volatile missions, while balancing the time and cost efficiency of these learning experiences. Hence, across coalition nations, many defense institutions are increasing their investments in innovative learning science and technologies (Raybourn, Schatz, et al., 2017), even while they may be reducing investments in more traditional delivery modalities. Some innovations, such as lifelong Artificial Intelligence (AI) driven personalization, require more research; however, other capabilities represent viable near-term, low-cost, high-reward solutions.

Blended learning (e.g., ADL combined with live training) represents one of these near-term opportunities. Backed by research showing that it can improve training effectiveness and efficiency (e.g., Fautua, Schatz, et al., 2014; U.S. Department of Education, 2010), and promoted by the U.S. Department of Defense Instruction 1322.26 ("Distributed Learning"), blended learning promises to be a high-value investment. Consequently, implementing blended learning in training exercises is an obvious and efficient mechanism to enhance those high-investment events.

The U.S. Joint Staff J7's Blended Learning–Training Program demonstrated that e-learning, web-based small group scenarios, and associated pre-training metrics could be advantageously blended into training exercises for the U.S. Combatant Commands (Fautua, Schatz, et al., 2014). However, large multinational training exercises, including CAX events, are generally conceived and delivered separately from ADL-based activities. Moreover, even in the U.S. Blended Learning–Training Program, trainee data collected from ADL-based activities is organizationally siloed across services, interagency, and partner nations. ADL-based performance, if meaningfully assessed at all, is not readily comparable to the exercise outcomes.

Related assessment issues affect exercises, too. For instance, exercise performance is rarely assessed quantitatively against strategic objectives, and the performance assessments that are conducted typically do not provide relevant,

comparable data on individual trainees' performance. In short, while we may know if the training was completed, it is more difficult to objectively answer the basic question: did the exercise do any good?

xAPI, developed by the U.S. ADL Initiative, is an emerging technical standard that supports learning analytics. Although xAPI has gained popularity and small-n field tests have delivered encouraging early results (Gallagher, et al. 2017), it has remained largely untested as a solution for delivering comparable results across complex multi-platform asynchronous learning and performance data feeds at scale. Organizers of the Viking 18 multinational exercise decided to run a proof-of-concept test on making ADL an integral part of this large event, including the use of xAPI to support trainee performance assessments across the ADL-based activities.

The Viking Exercise

The Viking exercise series was first chartered in 1999 as a Swedish and U.S. initiative at NATO's 50th Anniversary Summit. Since that time, the Swedish Armed Forces and Folke Bernadotte Academy have hosted Viking eight times. It has become the largest recurring civil-military relations exercise worldwide, with 61 countries and 80 organizations participating in 2018 (Swedish Armed Forces, 2018). The size and complexity of Viking make it an appealing case study. Its reliable cycles of iteration every three years also offer a foundation for regular retesting on lessons learned across the exercise series (Ljung, Ax, et al., 2018).

Viking 18 was held for ten-days in April 2018, across networked sites located in Brazil, Bulgaria, Finland, Ireland, Serbia and Sweden. It involved about 2,500 people, including 1,300 trainees and additional operators, monitors and support staff. The exercise trains civilians, military, police, and nongovernmental organizations together so that they are better prepared for deployment to an imaginary crisis response mission. The exercise scenario is multidimensional, multifunctional and multinational.

The ADL Effort in Viking 18: A Hard Case

Viking 18 marked the first example of full operational integration of ADL and xAPI learning analytics into a large scale, multinational collective training event. Viking organizers sought to approach the 2018 exercise as a total learning experience, incorporating more than two-dozen xAPI-enabled e-learning courses and a synchronous CAX simulation as well as learning analytics for both individual and aggregated performance in the e-learning courses and for the exercise objectives.

While the technical solutions for this initiative are increasingly mature, the complexity of multiple stakeholders matched with legacy platforms and stovepiped data, offered up an inherently challenging environment. That was exactly what we hoped for: a hard case, where countervailing conditions made success highly unlikely. Delivering a solution that worked would suggest that our model for operational integration of ADL into exercises has generalizable promise (Rapport, 2015).

The risks faced were both human and technical. The Learning Management System (LMS) used by Viking is a national solution that is not xAPI compliant. Much of the e-learning course content was provided by diverse organizations and was built using various standards, formats and technologies. Different national stakeholders used multiple exercise platforms, and there had previously been no attempt to integrate the data across the various exercise systems. The challenging cyber security environment added to the task complexity, both in the form of active threats from known real-world adversaries and in the form of local security measures enforced by participant countries, limiting access to online learning content. Implementation of the ADL integration proceeded on a limited time frame of less than twelve months with only three months allocated for technical effort. And finally, the Viking 18 exercise was a high visibility event, with significant presence by senior management and flag officers. While successful innovation often demands that we embrace the risk of failure, few are eager to fail in full view of executive leadership (Farson & Keyes, 2002).

VIKING 18 IMPLEMENTATION RESULTS

The integration of ADL into the Viking 18 exercise was a successful proof-of-concept in terms of cyber security, e-learning content interoperability, xAPI integration, user demand, and technology for data integration and visualization.

Cyber Security

The Viking 18 xAPI Learning Record Store and the visualization dashboard were hosted on a dedicated server, protected against Distributed Denial of Service (DDoS) and other forms of brute force attacks on the subnet level. The server was accessible only via a connection with strong encryption. All unused ports were disabled and direct server access was tightly restricted. Application level access was protected by mandatory strong passwords and failed login attempts control. All content was regularly backed up. The Secure Socket Layer (SSL) configuration was set to provide a good balance between security and compatibility with current browsers, including TLS v1.2 protocol, 2048-Bit RSA / 256 bit ECDSA Private Keys, (ECDSA first), and AEAD cipher suites.[1] All server activity was continuously logged and monitored for potential threats. These ADL systems remained live, with no intrusions, throughout the Viking 18 pre-training and exercise execution periods.

e-Learning Content Interoperability

A total of 29 e-learning courses were selected by the Viking 18 ADL Working Group based on their alignment with exercise objectives. Two courses were purpose-built, while the rest were contributed by six nations and NATO. All courses were in English. Participants could access the courses through the Swedish LMS one month prior to the exercise, and they were accessible throughout the event.

The Viking 18 ADL Working Group built a "course plan" for each Viking 18 exercise unit, which roughly aligned different e-learning objectives with training audience exercise roles. The course plans divided the e-learning courses into three categories. The Level 1 (L1) course, *Introduction to Viking*, was initially mandatory for all exercise participants but later downgraded to "highly recommended" after non-defense civilian organizers objected to utilizing a military platform for their e-learning. The L1 course familiarized participants with the organization of the exercise, the basic scenario and road to crisis narrative, so that the training audience might more successfully navigate the exercise. The set of L2 courses included contributions from of Joint Staff J7's Joint Knowledge Online, NATO Allied Command Transformation, and member nations from the Regional ADL Initiative (RADLI), an ADL defense cooperative in southeastern Europe including Serbia, Slovenia, Bosnia and Herzegovina, and Macedonia. L2 courses included topics aligned to the key exercise objectives, such as *Law of Armed Conflict* and *Protection of Civilians*. Two to three L2 courses, with e-learning content relevant for participants' specific roles, were assigned as "recommended" to each exercise unit. Finally, the L3 category of courses included all other e-learning courses of general interest to the Viking 18 training audience.

xAPI Integration

Viking 18 was the first large-scale test for xAPI in a multinational exercise. xAPI is a technical specification, developed by the U.S. ADL Initiative, that makes it possible to record, aggregate and analyze learning performance data across a multitude of diverse learning platforms and experiences. Despite the maturity of xAPI, integrating it into the multifaceted Viking experience, retrofitting it into the legacy SCORM[2]-based courseware, and instantiating the secure xAPI Learning Record Store (i.e., the repository for the xAPI-based data) proved challenging.

The Swedish Viking 18 LMS, *itslearning*, was not xAPI-compatible, meaning it lacked the capability to natively convert e-learning SCORM data to xAPI statements and store them in a local Learning Record Store. Thus, it was necessary to build xAPI functionality into the courses and to have the courses communicate directly with an external Learning Record Store. Integration of xAPI into the SCORM-based courses was accomplished utilizing the "xAPI Wrapper" and "SCORM-to-xAPI-Wrapper," a set of open-source JavaScript libraries that convert SCORM run-time data into xAPI statements (developed by the U.S. ADL Initiative; see github.com/adlnet/xAPIWrapper and github.com/adlnet/SCORM-to-xAPI-Wrapper).

The diversity of the 29 courses presented a real challenge for integrating xAPI into each of them. They were created by many different authoring tools, exported in at least seven different SCORM versions and often deployed on legacy platforms. The xAPI Wrapper requires minimal changes to instrument xAPI onto most newer SCORM courses, and

---

[1] The server Secure Sockets Layer configuration was tested by the SSLLabs Server Test and received the A grade: www.ssllabs.com/ssltest/analyze.html?d=vk18. jeffersonhosting.org
[2] Shareable Content Object Reference Model (SCORM)

its functionality can be replicated across a wide variety of legacy course content with some adjustment. However, the e-learning team had to create specific solutions on a case-by-case basis to implement the functionality on older courses. This sometimes required substantial modifications to both the SCORM-to-xAPI-Wrapper library and to the courses themselves. For instance, in some courses all SCORM communication was controlled by a single SCORM driver JavaScript file; in those cases, modification of the one wrapper file provided xAPI integration. However, in other cases, the courses were built as multiple learning chapters, and each chapter had to be wrapped separately. This was sometimes further complicated by each chapter containing multiple iFrames, with one responsible for the main SCORM communication and another for rendering course content. With content loading controlled by the *progress.js* file, the main problem was that all of these pages included the same *APIWrapper.js* file. Because some SCORM calls go from the parent *wrapper.html* file and some from the child *scorm.html* file, we had to avoid making several instances of the same wrapper object and at the same time control access to the parent frame object from the child frame, which is often blocked by browser security settings. Nevertheless, we successfully demonstrated that legacy courseware is ultimately no barrier to xAPI implementation.

Integration of xAPI Wrapper functionality across these 29 diverse courses enabled the collection of nearly any semantic data on learners' experiences in the courses; however, for the purposes of this proof-of-concept field test, we collected only course initiation, course completion, and test score data. Results were sent directly from the xAPI-enabled course into the cloud-based Learning Record Store, which resided on a separate server from the LMS.

High User Demand Despite Obstacles to Use

Organizational and usability obstacles created access barriers to the e-learning. Organizationally, there was little-to-no advertising of the e-learning resources to the training audience, and this lack of communication was reinforced when the *Intro to Viking* course was downgraded from "mandatory" to "highly recommended." From a usability perspective, access to the LMS proved cumbersome, at best. Participants were required to create multiple user accounts, and in many nations, Internet access at military facilities was so tightly controlled that participants opted to complete their e-learning after work hours.

Nevertheless, we saw a relatively high demand for the e-learning resources. Over 770 exercises participants—close to 40% of the training audience—successfully created e-learning accounts for the V18 LMS. More than 700 course-completions were recorded, and over 1,000 additional courses were initiated but not completed. Top courses completed, in rank order were: *Intro to Viking*, *Gender Awareness*, *SitaWare* (a primer on the exercise battle management software), *Exonaut* (a primer on the exercise simulator), *United Nations Peacekeeping Operations (UNPKO)*, and *Humanitarian Law*. Of these, three were authored with RADLI countries, one by U.S. Joint Knowledge Online, and two by a team lead by one of this paper's authors.

The difference between an initiated course and a completed course may vary due in part to how the courses were designed. Some of the courses only report "completed" if the learner moves through all the material, while some require the learner to complete a number of self-tests before reporting completion.
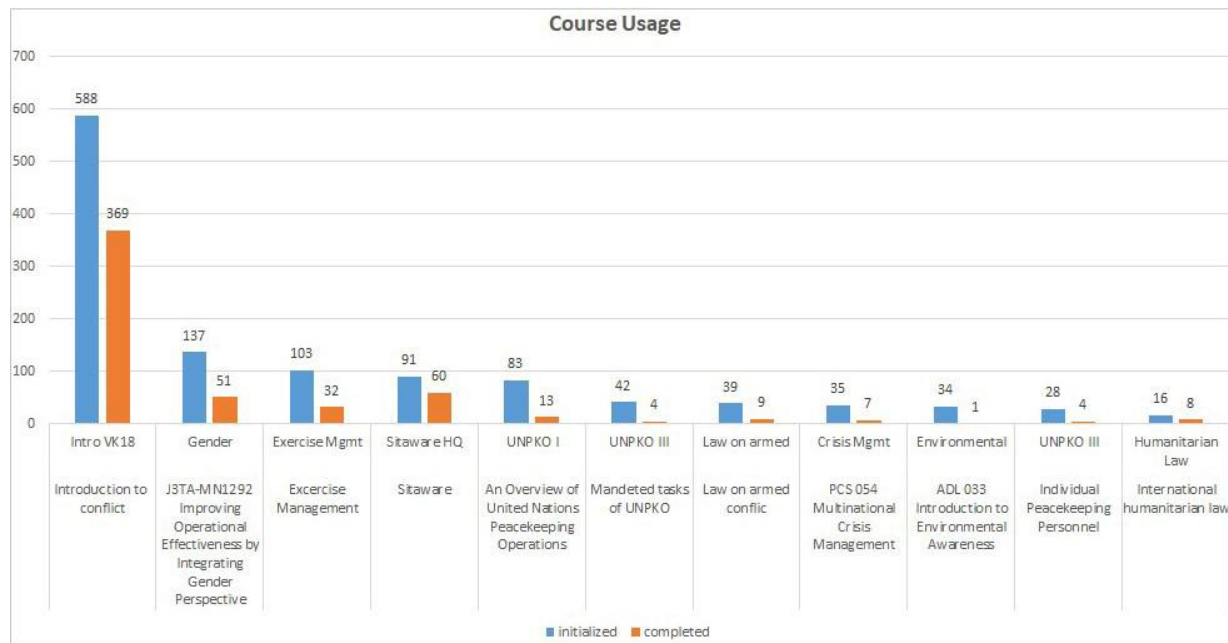
Figure 1. Top e-Learning Course Initiations and Completions

Data Integration and Visualization

To support analysis of the accumulated data, the ADL development team created a platform-neutral, web-based dashboard capable of managing asynchronous data streams from diverse sources, in multiple formats and at various scales. For Viking 18, the dashboard aggregated and visualized live streams of both xAPI-conformant data from the e-learning courses and post-hoc non-xAPI data from the exercise management and evaluation system. This meant that exercise organizers, participants, and other stakeholders could trace individuals' performance across time and platforms, unlocking the potential for deeper insights into training outcomes before, during and after the exercise.

The dashboard also produced "Nation Pages," summarizing the aggregate performance of participants from a given national unit. These proved highly popular. The Nation Pages offered a useful summary for participant delegations: something clear, immediate and relevant to bring home. Finland and Sweden were among the top performers in terms of e-learning course completions, scores, and diversity of courses taken.

The Viking 18 learning analytics dashboard demonstrates that the technology and knowhow is available to support the integration and visualization of data across multiple systems, to meet the aspirations of higher-level data-driven learning and to provide actionable insights to learners; observers, mentors and trainers; operators; planners; and senior leaders (Lang et al., 2017).

The following pages show several screen-captures from the Viking 18 dashboard (refer to Figures 1–5).
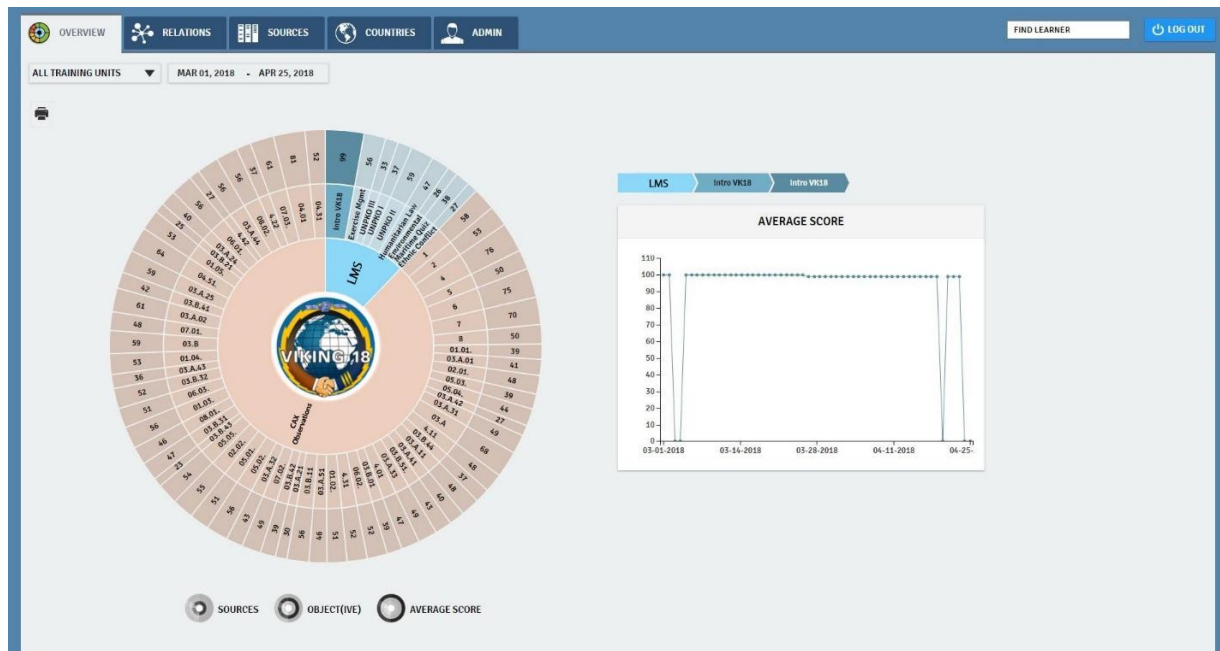
Figure 2a. Dashboard: Home

The dashboard homepage offered summary charts on top learners, top learning activities, and top learning objects, as well as an interactive radius graph displaying time series data on each e-learning course and on each exercise objective. It highlights the ability to simultaneously integrate and visualize data from multiple platforms.
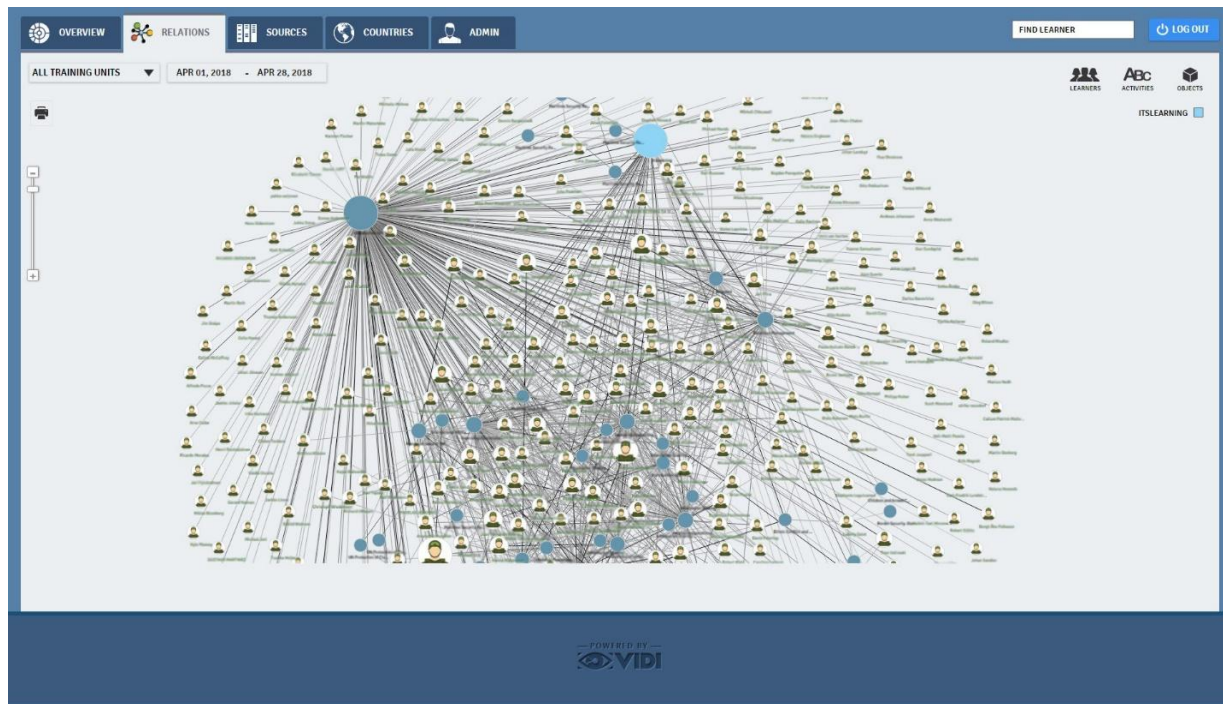

Figure 2b. Dashboard: Home

Figure 3. Dashboard: User Relations

While top-down structure is important, social networks and informal organic clustering is inevitable and important for understanding the social elements of a learning environment. The Viking 18 dashboard captures this by way of a network diagram showing clusters of students and courses they have taken. (In the real dashboard, exercise organizers can see relevant details about the individuals, but that personally identifiable information is removed from the images in this article.)
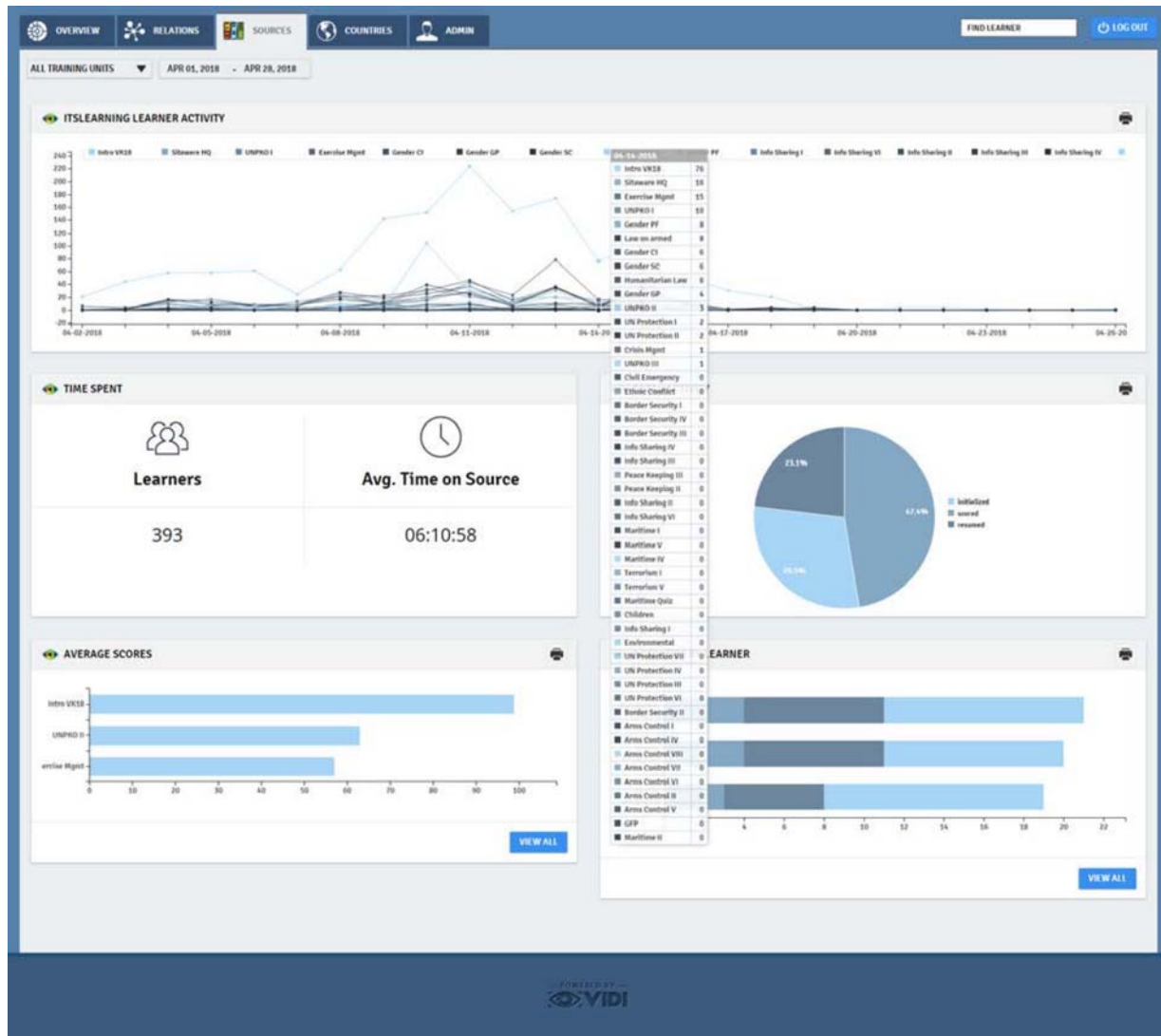
Figure 4. Dashboard: e-Learning Detail

The dedicated page for e-learning data displayed time series results for all courses in one view, average scores on top courses, and summary performance for top learners. Not surprisingly, the time series data revealed a surge of e-learning activity starting two weeks before the exercise, continuing into the third day of the event and then tapering off to occasional use by the fifth day of the ten-day event.
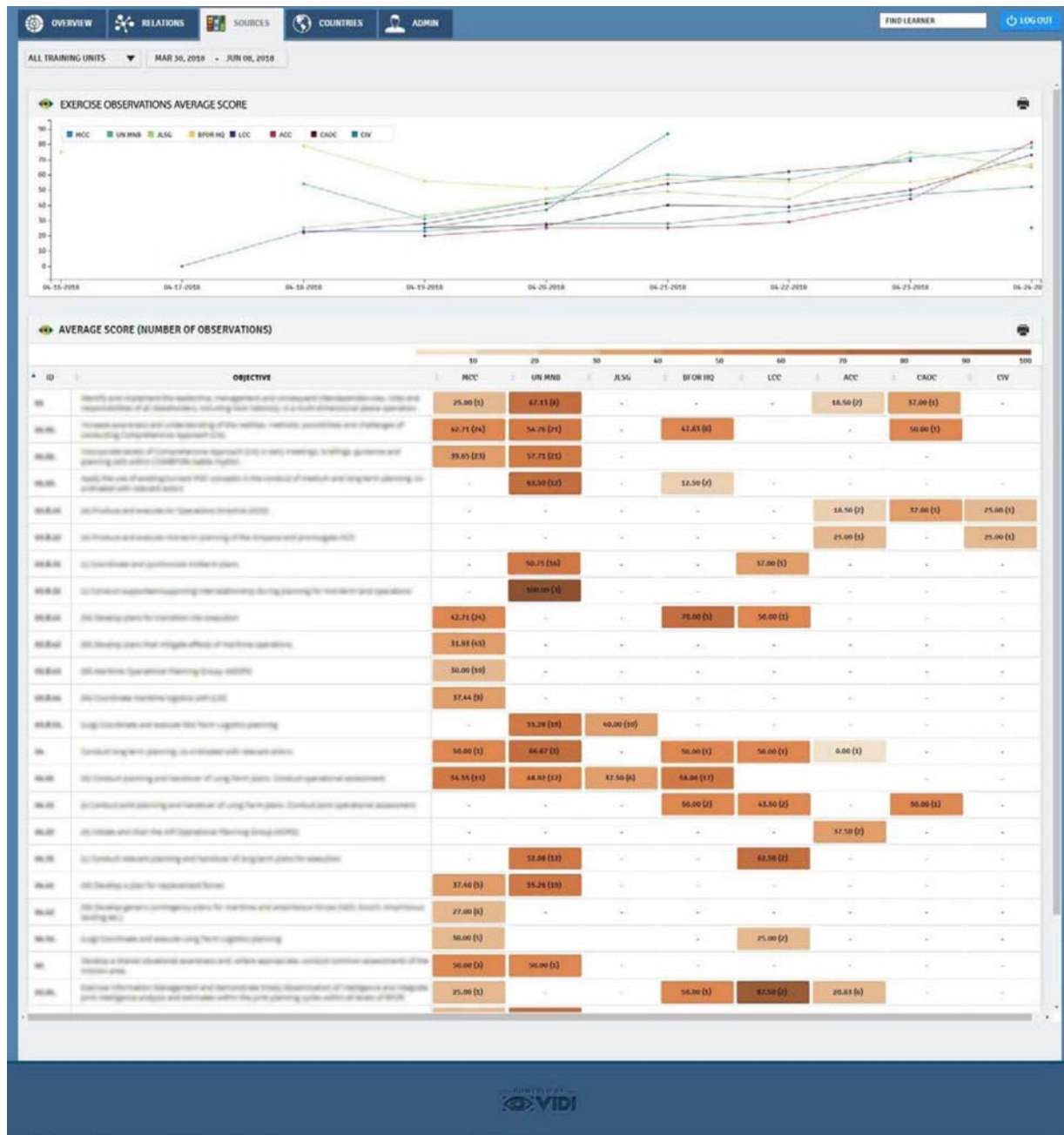
Figure 5. Dashboard: Average Exercise Observation Scores by Group (randomized data depicted for illustration – real data was shown to exercise participants)

The dedicated detail page for exercise observation scores summarized both time series and aggregate scores by objective and command unit. The timeline visualization shows a clear upward trend in exercise observation scores.
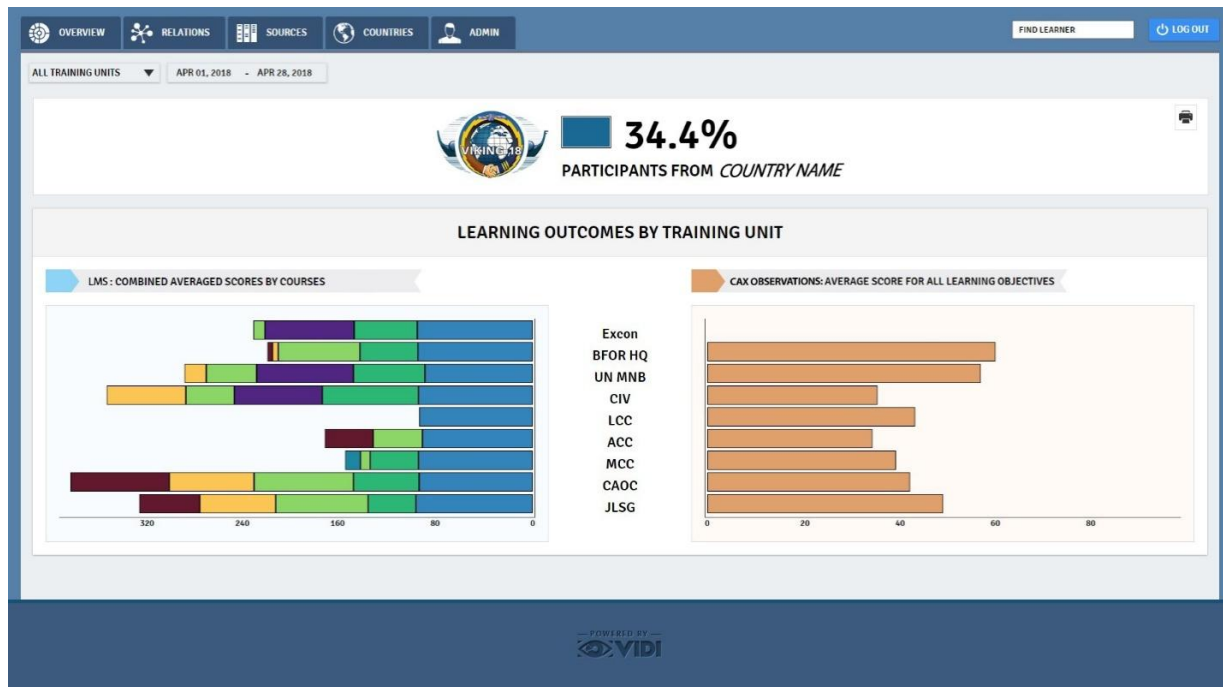
Figure 6. Dashboard: National Summary - Unit Results for Courses and Exercise Observation Data
(depicts an anonymized example)

The national results summary page on the dashboard offered a quick take-away for national contingents. On the left side is sum of the average scores on courses by users from one training unit - each course has a different color (LMS data). On the right side is the average observation scores for all learning objectives by those same training units.

ANALYSIS

According to the final Viking 18 manning list, the training audience included 1303 participants, including 376 off-site personnel. Other individuals took part in the exercise (e.g., planning, evaluation and general logistics personnel and support technicians, exercise drivers and operational staff) but were not evaluated in the exercise management platform. We collected e-learning data on the 773 participants with LMS accounts and on all participants evaluated in the exercise itself, but at the command unit or team level. Of the 773 participants with LMS accounts, 608 were in the training audience, while 165 LMS accounts were created by members of Exercise Control team.

The training audience was divided into eight main units: (1) BFOR HQ, the NATO Crisis Response Operation in the fictional Bogaland; (2) MCC, Maritime Component Command; (3) ACC, Air Component Command; (4) JLSG, Joint Logistics Support Group; (5) LCC, Land Component Command; (6) UN MNB, the United Nation Mission in the fictional Bogaland; (7) CAOC, the Combined Air Operations Centre; and (8) CIV, the other civilian organizations.

While our sample size is small and does not allow for statistical analysis, we can make some initial observations on the relationship between e-learning and exercise performance by looking at the percentage of completions for the *Introduction to Viking* course by members of the training audience, relative to their exercise performance observation average scores, by command unit.

Figure 7 depicts the Introduction to Viking course completion rate compared to in-exercise performance scores, as collected from the exercise management and evaluation system. We excluded two command groups, which were outliers in terms of exercise performance observations, UN MNB and BFOR HQ, and so were at the same time least relevant for judging the effectiveness of the *Introduction to Viking* course. UN MNB included military personnel from all participating countries, many of whom were not fluent in English and likely less harshly judged by exercise

evaluators; while BFOR HQ participants were not only role playing in the exercise, but actually served in HQ units in their regular jobs, so orientation was less relevant for their success in the training.
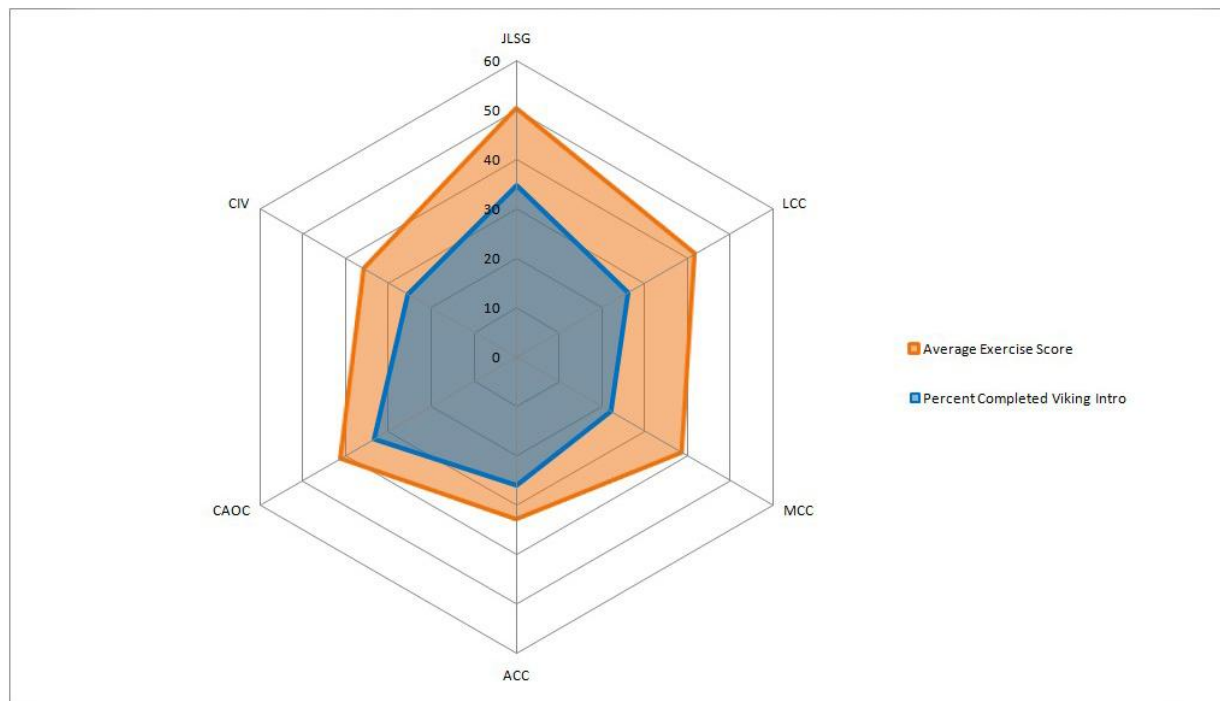

Figure 7. Viking 18 Introduction Course Completion vs. Average Exercise Score by Command Group

The radar graph in Figure 7 clearly suggests a pattern of improved performance in exercise observation scores with higher rates of Introduction to Viking course completion by the MCC, LCC, JLSG and CAOC units. The CIV unit, which was flagged off the *Introduction to Viking* course by their organizational leadership, significantly underperformed on exercise observations. On the other hand, ACC breaks the pattern. They showed a high rate of completion on the *Introduction to Viking* course, but finished with the lowest exercise observation scores. All together, there is a great deal of room here for further research.

CONCLUSION AND LESSONS IDENTIFIED

The integration of ADL and xAPI into Viking 18 was a success, particularly as an initial capability demonstration, with many lessons for future implementations of blended learning in multinational exercises.

- E-learning course objectives and exercise objectives should better align.
- The focus of learning analytics should be driven by the demand signals of stakeholders: learner, operational, and strategic.
- User access to both learning content and to learning analytics must be improved.

Viking 18 demonstrated the viability of blending ADL into multinational exercises, integrating xAPI across diverse e-learning courseware, extracting xAPI from a non-compliant LMS, executing learning analytics at a large scale and visualizing disparate types of data, in real time, within a multinational training context. Viking 18, with its scale, diversity, legacy platforms, cyber security challenges and accelerated ADL development schedule proved the "hard case." If xAPI and learning analytics can be successfully deployed in this context, they can be used anywhere.

The Viking 18 learning analytics and data dashboard show the power of even simple frequency data to glean deeper insights from exercise training outcomes for data-driven learning. This initial proof-of-concept demonstrates how much can be achieved, with lightweight ADL capabilities and close collaboration with coalition partners.

ACKNOWLEDGEMENTS

REFERENCES

Farson, R., and Keyes, R. (2002). *Whoever makes the most mistakes wins: The paradox of innovation*, New York: Free Press.

Fautua, D. T., Schatz, S., Reitz, E., & Bockelman, P. (2014). Institutionalizing blended learning into joint training: A case study and 10 recommendations. In *Proceedings of the I/ITSEC*, Arlington, VA: NTSA.

Gallagher, P. S., Folsom-Kovarik, J.T., Schatz, S., Barr, A., Turkaly, S. (2017). Total Learning Architecture Development: A Design-Based Research Approach. In *Proceedings of the I/ITSEC*, Arlington, VA: NTSA.

Lang, C., Siemens, G., Wise, A., Gašević, D. (2017). *Handbook of learning analytics.* Society for Research on Learning Analytics (SoLAR). Retrieved from solaresearch.org/hla-17/.

Ljung, N. Ax, T., Presnall, A., & Schatz, S. (2018). Integrating Advanced Distributed Learning into Multinational Exercises, *Proceedings of the* I/ITSEC, Arlington, VA: NTSA.

Rapport, A. (2015). Hard thinking about hard and easy cases in security studies. *Security Studies, 24*:3, pp. 431-455.

Raybourn, E.M., Schatz, S., Vogel-Walcutt, J., & Vierling, K. (2017). At the tipping point: Learning science and technology as key strategic enablers for the future of defense and security. *Proceedings of the I/ITSEC*, Arlington, VA: NTSA.

Swedish Armed Forces (2018). Viking 18. Blog post. Retrieved: https://blogg.forsvarsmakten.se/viking18/

Schatz, S., Fautua, D., Stodd, J. & Reitz, E. (2015). The changing face of military learning. *Proceedings of the I/ITSEC*, Arlington, VA: NTSA.

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development (2010). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies.* Washington, D.C.