

Interoperable assessments using HPML: A novice conning skills acquisition use case.

Bruno Emond
National Research Council Canada
Ottawa, Ontario
Bruno.Emond@nrc-cnrc.gc.ca

ABSTRACT

Demands in training require organizations to maximize access to adaptive training through high and low-fidelity simulations. However, an increase of learning opportunities through simulations must also be associated with a high level of efficiency and efficacy of the training system as a whole (Atkinson & Killilea, 2015). To avoid the risk of creating a collection of practice simulators instead of objective-based training environments (Stacy, Merket, Freeman, Wiese, & Jackson, 2005), training simulators must provide a precise yet comprehensible means to express and manipulate measurements, and assessments across a range of learning opportunities (Stacy, Ayers, Freeman, & Haimson, 2006). The motivation for the work reported in this paper originated from the need to redeploy measure and assessment software from a training simulation application to another. Given our interest to repurpose investments in learning design, the main research questions the current paper seeks to address consist of determining: 1) what approach would best fit interoperable measure and assessment computations, and, 2) to what extent the selected approach is adequate to represent specific measures and assessments we had implemented in our training simulation. The first section briefly presents major interoperable assessment initiatives. The section concludes that the Human Performance Markup Language (HPML) seems to best fit our interoperable measure and assessment needs, which is to repurpose and allow interoperability of measure and assessment computations. The Human Performance Markup Language (HPML) aims at fulfilling this purpose by providing a simple and reusable way to represent the performance of individuals and teams in those systems (Walker, Tolland, & Stacy, 2015). HPML supports the representation of measurements and assessments, and how they relate to performance and learning data, as well as training objectives. In the latter case, the HPML training objective package for instance, provides a scalable formal mechanism to document and manage training objectives, their relationships to scenario conditions, and performance measures (Stacy & Freeman, 2016). The second section gives an overview of HPML, followed by a presentation of a target use case, a training simulation for novice ship conning skill acquisition. The third section discusses how some HPML assessment templates can be applied to the use case. The application of HPML to the use case indicated that most of the assessment computations that were used in the training simulation for novice ship conning skill acquisition could be represented. A possible extension to HPML for expressing otherwise cases in category selection was identified, which would simplify assessment templates. However, the sparse HPML documentation, and low number of examples available made it difficult at times to determine if the analysis of our use case respected the intention of the HPML standard proposal. In this respect, future work is needed to evaluate how expressive is HPML and what are the HPML limits and boundaries in the learning technology value chain.

ABOUT THE AUTHORS

Dr. Bruno Emond is a senior research officer at the National Research Council Canada. He joined NRC in 2001 and holds a B.A. and M.A. in philosophy, and a Ph.D. in educational psychology from McGill University. His research evolved over his career on issues related to knowledge representation, logic, text comprehension, and cognitive modelling. Dr. Emond's current interests focus on adaptive training systems, and educational data mining.

Interoperable assessments using HPML: A novice conning skills acquisition use case.

Bruno Emond
National Research Council Canada
Ottawa, Ontario
Bruno.Emond@nrc-cnrc.gc.ca

INTRODUCTION

The motivation for the work reported in this paper originated from the need to redeploy measure and assessment software from a training simulation application to another. A fair amount of efforts was put into the original training application measure and assessment components, and moving to another platform involved a significant risk of losing investments in learning design, when the performance evaluation logic was embedded (locked) in the training simulation application code. Our previous work (Emond et al., 2016) used the xAPI framework, but even though xAPI provides the capability to repurpose recorded measure and assessment data, it does not provide the means to repurpose the computations that generated this data. As long as the application development is carried within the same development environment, software reuse is relatively easy. However, this is not always possible in a fast-paced market of technological innovations relying on a commercial-off-the-shelf software integration model, and networked software services. Given our interest to repurpose investments in learning design, the main research questions the current paper seeks to address consist of determining: 1) what approach would best fit interoperable measure and assessment computations, and, 2) to what extent the selected approach is adequate to represent specific measures and assessments we had implemented in our training simulation.

The paper discusses interoperable performance assessments. The first section briefly presents major interoperable assessment initiatives, including CASS, IMS Caliper, HPML, and xAPI. The section concludes that HPML seems to best fit our interoperable measure and assessment needs, which is to repurpose and allow interoperability of measure and assessment computations. The second section gives an overview of HPML, followed by a presentation of a target use case, a training simulation for novice ship conning skill acquisition. The third section discusses how some HPML assessment templates can be applied to the use case. Finally, the paper concludes by summarizing the main findings.

INTEROPERABLE ASSESSMENTS

Demands in training require organizations to maximize access to adaptive training through high and low-fidelity simulations. However, an increase of learning opportunities through simulations must also be associated with a high level of efficiency and efficacy of the training system as a whole (Atkinson & Killilea, 2015). To avoid the risk of creating a collection of practice simulators instead of objective-based training environments (Stacy et al., 2005), training simulators must provide a precise yet comprehensible means to express and manipulate measurements, and assessments across a range of learning opportunities (Stacy et al., 2006). In addition, given the rate of change in technology, it is very also important not to lock investments in learning design, measurements and assessments in specific customized training simulations. As new training technology emerges, being able to reuse software components for measures and assessments in a new training environment allows maintaining continuity and providing learning management refinement opportunities.

Decoupling of software components has been a key driver in commercial-off-the-shelf software integration, web services, and cloud computing. Component-based software development proposes to accelerate software development and reduces development costs by assembling systems from existing components. Software components are modules that encapsulates related data and functions, software packages or web services. Components interact as well as use the services of each other through interfaces. Inner functionality or the structure of the components is encapsulated or is not known to the component user (Gupta, 2012). Components based development includes option of purchasing components from the market rather than building the components from scratch, or if the market does not offer what is needed (Brownsword, Oberndorf, & Sledge, 2000). The training

simulation market has been subject to these trends and opportunities as well, especially given the increased availability and rate of change in game engine platforms and virtual reality devices, adding pressure to provide assessments interoperability across learning platforms.

There are currently various efforts to develop technology to support interoperable performance assessments. Major initiatives include CASS / Competency And Skill System (CASS Project, 2016), IMS Caliper Analytics (IMS Global Learning Consortium, 2015), Total Learning Architecture (Folsom-Kovarik & Raybourn, 2016), and xAPI (Advanced Distributed Learning & ADLnet.gov, 2017). The xAPI framework has received a fair amount of attention (Goodwin, Murphy, & Medford, 2016; Hruska, Long, & Amburn, 2014; Poeppelman et al., 2013) and is a key component of the Total Learning Architecture (Folsom-Kovarik & Raybourn, 2016). The main objective of xAPI is to offer a simple format for encoding learning events, as well as requirements for a learning record store independent of learning and training applications. The xAPI framework proposes a simple format to encode learning events and results in the form of time stamped “actor-verb-object” statements, with additional keys like “result”, to add details to a measured outcome, or “context”, which can contain relevant information to contextualize learning event statements. In xAPI, learning events are stored in a learning record store (Advanced Distributed Learning, 2017). The focus of xAPI is mostly on the encoding and storage of learning events, leaving the details unspecified on how measures, and assessments were computed to generate the recorded results.

The Total Learning Architecture (Folsom-Kovarik & Raybourn, 2016) seeks interoperability by the separation of learning record providers (training simulations), learning record consumers (learning analytics), and learning record stores (learning event storage). CASS aims to facilitate the use of competencies and competency frameworks by using an interoperable infrastructure between software systems for the production and consumption of competencies and competency assertions (CASS Project, 2016). IMS Caliper’s main goal is to establish a way to capture and obtain measures from a set of learning activities, each having one or several associated metric profiles defining the information flow from a learning activity.

Another less known approach to interoperable performance assessments is the Human Performance Markup Language (HPML), which aims at fulfilling the purpose of managing measurements, assessments, and training objectives by providing a simple and reusable way to represent the performance of individuals and teams in those systems (Walker et al., 2015). Other potential benefits of HPML include (Atkinson & Killilea, 2015): an increase in visibility and definition of measurement and assessment requirements, reduced acquisition lifecycle costs in the long term, and increased technical quality of human performance assessment. HPML is currently being examined as a potential standard by a Simulation Interoperability Standards Organization (SISO) study group (Simulation Interoperability Standards Organization, 2016a).

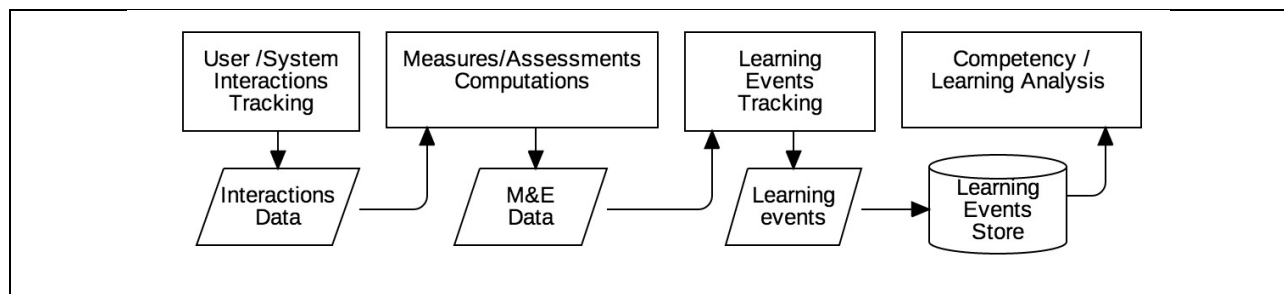


Figure 1. Data processing workflow from user/system interactions to competency modelling and learning analytics.

Figure 1 presents a data processing workflow from the capture of user/system interaction data to the analysis of learning, and the application of competency frameworks. As such, the figure is very general and ignores key elements related to the deployment of interoperable assessments, which in the current main stream efforts, happen on top of a shared learning events storage. However, Figure 1 helps to see that there are challenges to the deployment of interoperable assessments upstream from a learning events store. Current practice in learning events recording consists of capturing learning progress, and learning and performance outcomes, but not the specifications of how outcome values were computed. Table 1 presents a comparison between CASS, IMS Caliper and HPML as a function of the data processing workflow of Figure 1. The table identifies that user/system interactions are captured

at the application level. The tracking of learning events, as well as their further processing and analysis is also identified for CASS and IMS Caliper. However, only HPML offers means to represent explicitly how measures and assessments are performed. It appears then that HPML would best fit our interoperable measure and assessment needs, which is to repurpose and allow interoperability of the measure and assessment computations. The next section gives an overview of HPML.

Table 1. Comparison between CASS, IMS Caliper and HPML as a function of data processing workflow (Figure 1).

	User/System Interactions Tracking	Measure and Assessment Computations	Learning Events Tracking	Competency Modelling / Learning Analytics
CASS	Training/Learning Application	Training/Learning Application	xAPI statements & Learning Record Store	Competency frameworks & Learner profiles
IMS Caliper			Sensor API & Caliper Event Store	Metric profiles
HPML		Data requests, Measure and Assessment Schemas	Measure and Assessment Instances	

HPML

The Human Performance Modelling Language (HPML) is intended to cover all meaningful aspects of human performance measurements in various training and operational environments (Simulation Interoperability Standards Organization, 2016a). It is composed of schemas such as computations, measures, assessments, results, instances and periods. In the context of HPML, a computation represents an algorithm performing a mathematical, or logical operation. Measures represent the application of computations to data sources to produce measurements. Assessments are the assignment of measures to categories. Instances and periods represent the creation and use of measures and assessments for a given time and/or location context. Finally results refer to the output of both measure and assessment instantiations (Simulation Interoperability Standards Organization, 2016b). Figure 2 presents the dependencies between high-level HPML elements (A), and high-level XML schemas (B). The boxes added to the figure on the left represent a grouping of measure and assessment specifications or templates (red), and measure and assessment instances, and results as a function of specific performance data in time periods. The right-hand side of the figure shows a detailed list of all high-level HPML schema elements, including ones for representing data requests, constants, actors (individuals or teams), training objectives and their links to measures and assessments (training objective packages). Each of these top-level schemas can be expanded to cover many human performance-modelling requirements. As an example of the capability of HPML to human performance modelling requirements, Figure 3 gives a subset of HPML elements for the specification of measurement and assessment templates, which can be understood as specifications of how measures and assessments are calculated from various data sources (parameters, constants, and computations).

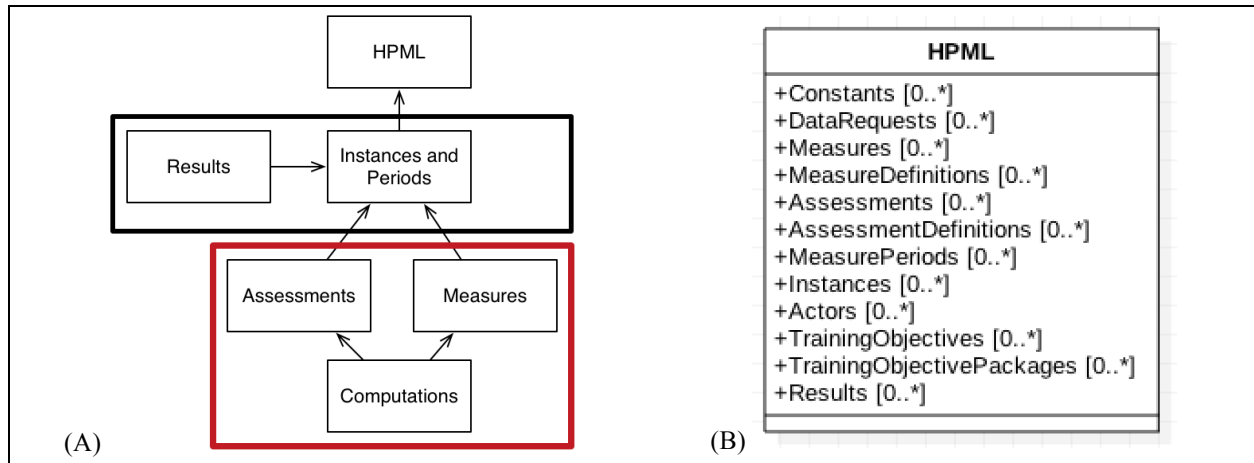


Figure 2. (A) High-level HPML abstraction, and (B) High-level XML schemas.

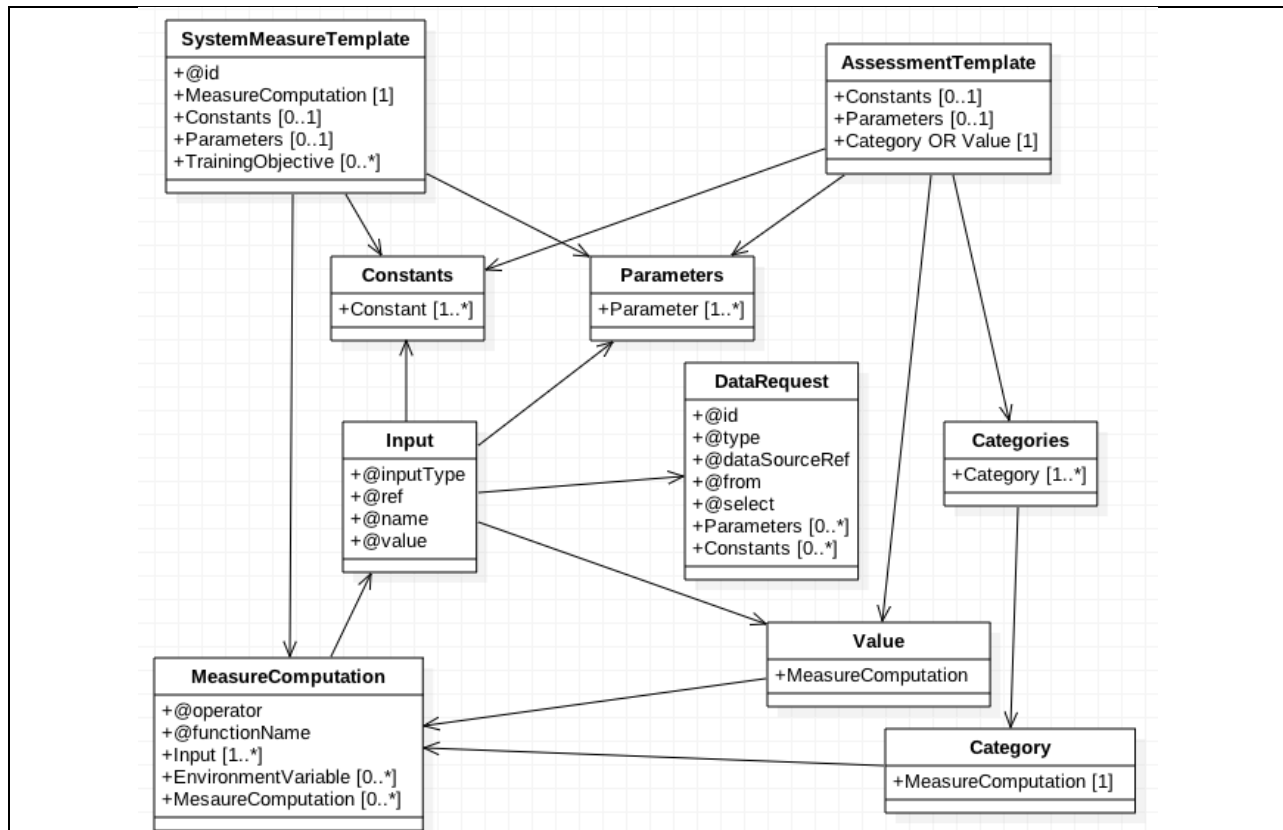


Figure 3. Subset of HPML elements for the specification of measurement and assessment templates.

HPML SCHEMAS FOR A TRAINING SIMULATION USE CASE

As stated in the introduction, the main research questions the current paper seeks to address consist of determining: 1) what approach would best fit interoperable measure and assessment computations, and, 2) to what extent the selected approach is adequate to represent specific measures and assessments we had implemented in our training simulation. The previous section identified HPML as a potential good response to the first question. The current section describes some of our efforts to apply the HPML specification to a training simulation for basic conning in order to answer the second question. The purpose of this mapping effort is to better understand the possibilities of HPML, using an existing training system as a use case, which is a simple training simulation for novices who have no prior knowledge or experience with ship conning. The general context for the development of this training

simulation (Emond et al., 2016), as well as cognitive models of novice learners in this simulation environment are described elsewhere (Emond & Vinson, 2017).

Simple Conning Training Tasks, and Training Objectives

The set of training tasks for the simulation was determined by looking at the range of possible simple course alteration commands given by a commanding officer to an officer of the watch. Simple course alteration commands were identified, and varied along two dimensions: the direction to turn the ship, and the new course heading. Both of these dimensions could be either explicitly stated or left implicit in the commanding officer request. Table 2 gives the three possible request cases, the combination of implicit direction and heading was left out because it does not allow the officer of the watch to determine a unique direction for the initial turn and final heading. The request by the commanding officer to “take the shortest path”, was added to force the trainee to think about which way the initial turn should be. This statement was added with a pedagogical interest given that trainees were encountering for the first time the calculation of optimal initial helm direction. From the three basic conning tasks in Table 2, two additional training tasks were designed. The first one aimed at simply allowing the trainee to master the association between “left-port”, “right-starboard”, and “straight helm-midships”. The other addition is an extension of the condition of explicit direction and implicit heading, where the ship must cross true north. For example, if the ship is initially on a course of 010 degree, a left turn of forty degrees will give a final heading of 330 degrees, adding some difficulty related to the calculation of the final heading requested. On the ship, headings are expressed in degrees (360), and north is both 000 and 360. Also, headings are expressed as a series of three digits, for example 125 is one-two-five, and not one hundred twenty-five.

Table 2. Simple course alteration orders by the commanding officer.

Direction	Heading	
	Explicit	Implicit
Explicit	Officer of the watch, come (right/left) to a new heading of (one-two-five).	Officer of the watch, come (right/left) (sixty) degrees.
Implicit	Officer of the watch, change course to a new heading of (one-two-five), taking the shortest path.	

In the training simulation, the learner who played an officer of the watch (OOW) interacted with a simulated captain, who issued requests verbally (text-to-speech software), and a simulated helmsman who simply executed the officer of the watch conning orders (speech recognition software). In the simulation, the helmsman does not repeat the conning orders, contrary to the navy protocol. Figure 4 presents the task workflow from the commanding officer request (orange), through the main three steps of a course alteration steps (blue), and the measurements and assessments (white) implemented as a means to track performance and learning. The measurements and assessments will be discussed in the HPML section below. Therefore, ignoring for the moment measures and assessments, upon hearing the commanding officer’s request, the first task for the officer of the watch is to determine which way the ship has to turn, and produce a conning order to reflect his/her choice (initial helm). Then the officer of the watch, depending on the state of the ship, could either issue a counter helm order to make small corrections as the ship approaches the new heading, or simply issue a steer order to steady the helm on the new course. However, depending on the success of the task sequences, the officer of the watch might be required to go through a new sequence of conning orders. Table 3 gives the list of simplified conning orders and conning protocol required to execute the training simulation course alterations.

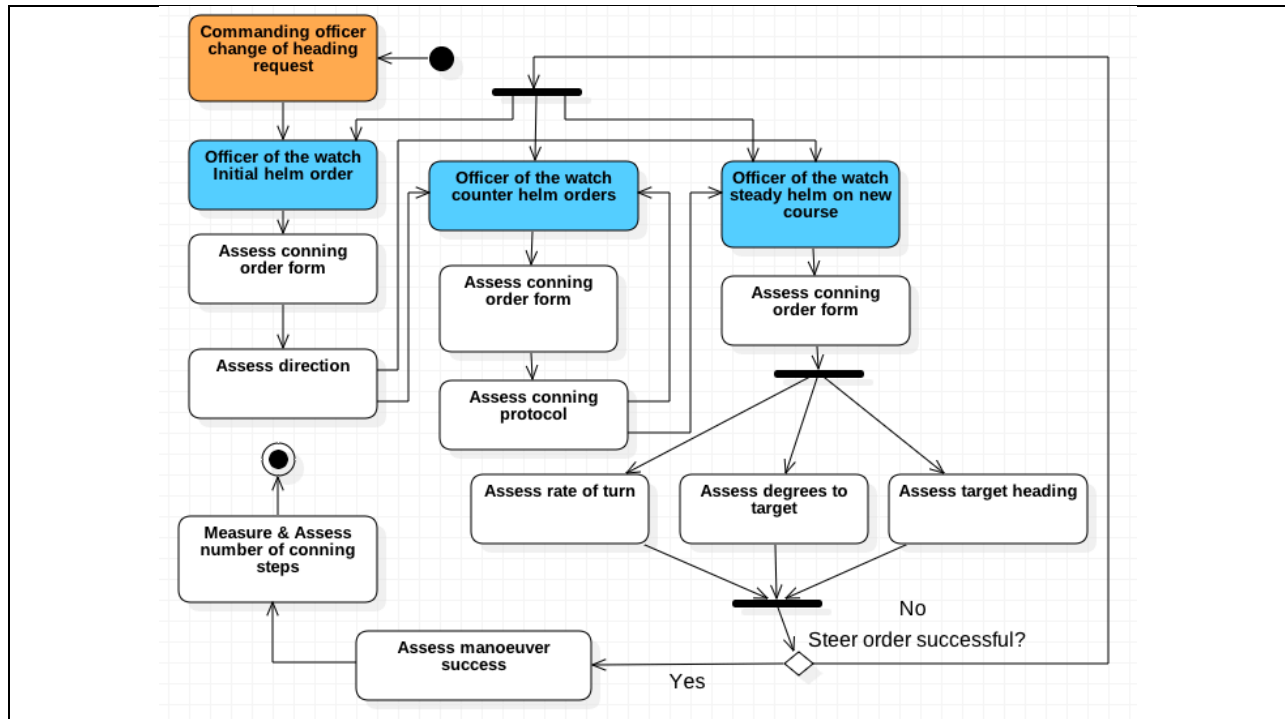


Figure 4. Basic conning training control flow: Commanding officer request (orange), Officer of the watch course alteration steps (blue), assessment and measurement (white).

Table 3. Simplified conning orders vocabulary for course alteration.

Course alteration steps	Examples	Interpretation
<i>Initial helm (helm order)</i> Direction + rudder angle	Port fifteen Starboard thirty	Turn left with a rudder angle of 15 degrees. Turn right with a rudder angle of 30 degrees.
<i>Counter helm</i> Midships + helm order	Midships, Port thirty Midships, Starboard fifteen	Bring the rudder to an angle of 0 degrees, then initiate a turn in the other direction.
<i>Steady helm on new course</i> Midships + steer order	Midships, Steer one zero five	Bring the rudder to an angle of 0 degrees, then order to steady the ship's direction on the new heading.

The basic conning training simulation contained five scenarios, which were designed to provide a basic introduction to conning orders for people having no prior knowledge in the area. The training objectives for each scenario were defined in terms of having evidence of mastering some conning order vocabulary, as well as planning and executing sequences of conning orders to execute a course alteration. Table 4 gives the training objectives for each of the scenarios. Given that we had no prior data to determine what would constitute an adequate definition of conning task mastery, we assumed that a certain number of consecutive successes in task execution could be used. Any failure in a task execution would reset the mastery count to zero. The mastery thresholds for each scenario are given in Table 4.

Table 4. Training scenarios (see Table 2 for examples of commanding officer requests)

Scenarios	Training objectives
Basic conning vocabulary. Port, Starboard, Midships	At the end of the scenario the learner will have shown evidence of mastering the basic conning orders using port and starboard at a standard rate of turn, as well as the midships conning order to stop the ship from turning. The mastery threshold was set at three consecutive successes each for port, starboard, and midships.
Explicit direction and explicit heading.	At the end of the scenario the learner will have shown evidence of mastering the basic conning orders from changing the direction of the ship, given the known initial turning direction and true bearing. The mastery threshold was set at two consecutive successes for a complete change of course direction.
Implicit direction and explicit heading	At the end of the scenario the learner will have shown evidence of mastering the basic conning orders from changing the direction of the ship, given the unknown initial turning direction and true bearing. The mastery threshold was set at two consecutive successes for a complete change of course direction.
Explicit direction and implicit heading.	At the end of the scenario the learner will have shown evidence of mastering the basic conning orders from changing the direction of the ship, given the known initial turning direction and relative bearing (unknown true bearing) using simple mental calculations. The mastery threshold was set at two consecutive successes for a complete change of course direction.
Explicit direction and implicit heading (crossing north).	At the end of the scenario the learner will have shown evidence of mastering the basic conning orders from changing the direction of the ship, given the known initial turning direction and relative bearing (unknown true bearing) using more demanding mental calculations. The mastery threshold was set at two consecutive successes for a complete change of course direction.

Application of HPML Schemas to the Conning Use Case

A complete application of HPML to the training application described in the previous section would cover many HPML schemas outlined in Figure 2 and Figure 3. The instantiation of HPML schemas can be fairly complex, as exemplified by Figure 3. The current section focuses only on the measurement and assessment specifications related to the conning training workflow of Figure 4, leaving aside issues of their specific applications to generate instances and results.

A frequent assessment in the basic conning use case is the assessment of the good form of simple conning orders. This assessment consists of determining from a speech recognition output if an utterance corresponds to the vocabulary and syntax of the simple conning options defined for a training simulation scenario. Figure 5 covers cases for the simplest scenario where the only acceptable utterances are “Port fifteen”, “Starboard fifteen”, and “Midships” (first row of Table 4). The tree structure represents how sub-elements of the assessment template are composed together. The first node under the assessment template groups the categories that will be used for the assessment, the categories are simply “success” and “failure”. Below each category are measure computation elements that need to be executed to determine which of the category should be the assessment outcome. The possibility of expressing computations in HPML is a very powerful feature that allows representing how measures and assessments are determined. From Figure 5, one can see that the success of the basic vocabulary assessment depends on the fact that the data source provided by the simulation matches one of the three possibilities. The second assessment category will always return the value “failure” if the first category does not evaluate to true. HPML assumes that the category evaluation proceeds from the first to the last in the group of categories. In Figure 5, the last category a workaround to handle “if everything else fails”. However, the categories list could contain an “otherwise case” category, such as `<Category type='Otherwise' name='Failure' />`. This HPML extension would simplify the assessment template.

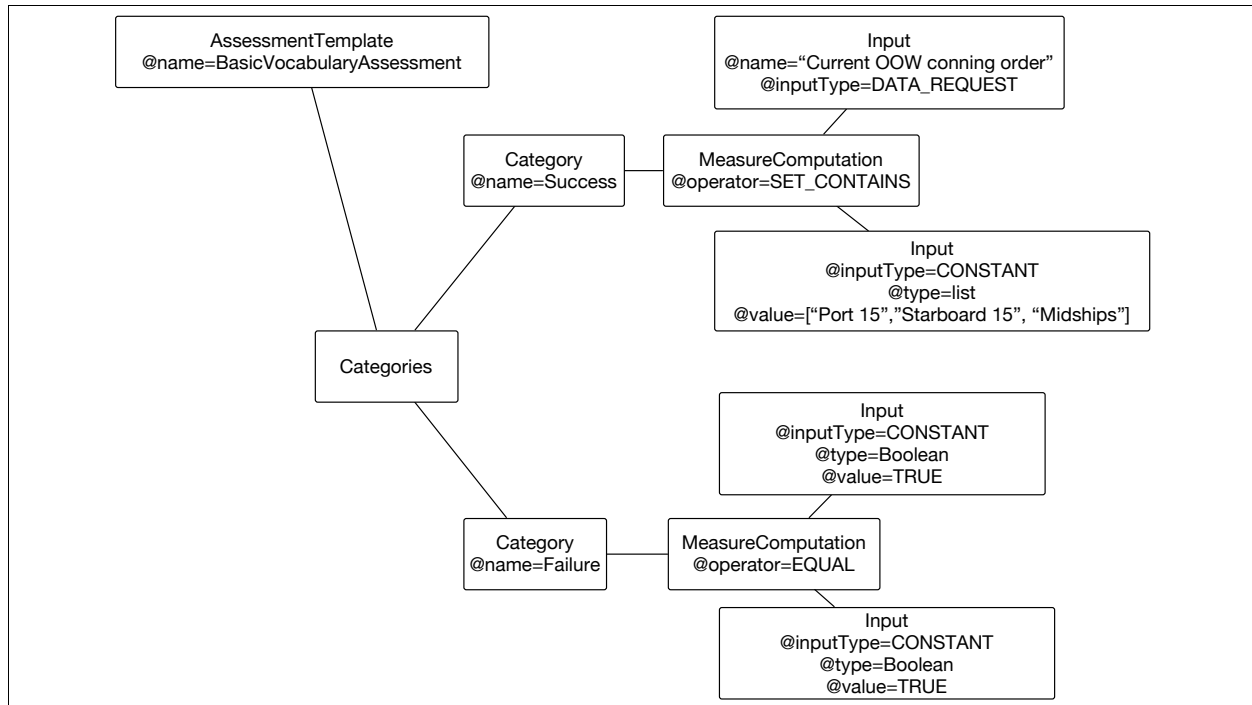


Figure 5. Form assessment of conning orders (basic vocabulary).

The assessment template defined for the evaluation of conning order forms is very similar to other assessments needed for the training simulation, including assessments of the initial helm direction, rate of turn, degrees to target, and steer heading exactness. In spite of differences in the number and type of measure computation, and comparing strings and/or numbers, the structure is identical with a category for “Success”, and another one for “Failure”. Table 5 identifies the main measure computation operators, data requests, constants, and parameters to cover all of the use case assessments. However, the reference to parameters holding values is speculative at this point in time, because the HPML specification does not offer an *ASSIGN_VALUE* operator that would allow setting parameter values. A second possibility would be to allow reference to uniquely named measurement instances holding values. A third possibility is to rely on the training application to keep track of key values, which could then be retrieved by data requests or by means of HPML environment variables. However, this last option combines measure and assessment computations with application development, which blurs the intended separation between learning design and software development.

Table 5. Assessments (Figure 4) required HPML measure computation operators and data requests.

Assessments		
	MeasureComputation operators	Datarequests(DR) / Constants(C) / Parameters(P)
Conning order form (Figure 5).		
	SET_CONTAINS, EQUAL	DR- Current OOW conning order C- List of well-formed conning orders
Conning orders sequence conformance (ex. Midships must occur between change of directions conning orders)		
	AND, OR, SET_CONTAINS, EQUAL	DR- Previous OOW conning order DR- Current OOW conning order C- List of orders
Initial change of direction (ex. "Port 15" when anticipated initial direction is left).		
	EQUAL	DR- Initial OOW conning order DR- Anticipated initial direction
Rate of turn		
	LESS_THAN	DR- Rate of turn at time of a Steer conning order C- Maximum acceptable rate of turn
Degrees to target		
	LESS_THAN	DR- Difference of degrees to target at the time of a Steer conning order C- Maximum acceptable difference in degrees
Exact new heading		
	EQUAL	DR- Heading value in the current Steer coning order DR- Requested new heading value by the Captain
Change of course success (there is no failure in any of the change of course steps).		
	NOT, SET_CONTAINS	P- List of change of course step assessments. C- Failure category
Change of course mastery		
	EQUAL	P- Count of successive change of course successes for a scenario C- Mastery count criterion

CONCLUSION

The main research questions consisted of determining: 1) what approach would best fit our interoperable measure and assessment needs, and, 2) to what extent the selected approach is adequate to represent our measure and assessment calculations. After reviewing some interoperable assessment initiatives, it appears that HPML would best fit our interoperable measure and assessment needs, which is to repurpose and allow interoperability of the measure and assessment computations. As to the second research question, the application of HPML to the use case indicated that most of the assessment computations that were used in the training simulation for novice ship conning skill acquisition could be represented. A possible extension to HPML for expressing otherwise cases in category selection was identified, which would simplify assessment templates. However, the sparse HPML documentation, and low number of examples available made it difficult at times to determine if the analysis of our use case respected the intention of the HPML standard proposal. In this respect, future work is needed to evaluate how expressive is HPML and what are the HPML limits and boundaries in the learning technology value chain.

ACKNOWLEDGEMENTS

I would like to thank Perry McDowell, Robby Robson, and LCdr Maxime Maugeais (Royal Canadian Navy) for their valuable comments to improve the quality of this paper.

REFERENCES

Advanced Distributed Learning. (2017). GitHub - adlnet/xAPI-Spec. Retrieved June 9, 2017, from

- <https://github.com/adlnet/xAPI-Spec>
Advanced Distributed Learning, & ADLnet.gov. (2017). xAPI. Retrieved May 19, 2017, from <https://www.adlnet.gov/xAPI>
- Atkinson, B. F. W., & Killilea, J. (2015). A review of the potential return on investment benefits of a human performance measurement standard: Lessons learned in the Navy aviation community. In *2015 Fall Simulation Interoperability Workshop, SIW 2015*. Orlando, FL: SISO - Simulation Interoperability Standards Organization.
- Brownsword, L., Oberndorf, T., & Sledge, C. A. (2000). Developing New Processes for COTS- Based Systems. *Software, IEEE, 17*(4), 48–55. <http://doi.org/10.1109/52.854068>
- CASS Project. (2016). CASS Documentation. Retrieved August 9, 2017, from <http://docs.cassproject.org/index.html>
- Emond, B., Maugeais, M., Vinson, N. G., Saikh, K., Fournier, H., Lapointe, J.-F., & Martin, C. (2016). Adaptive training simulation using speech interaction for training navy officers. In *Interservice/Industry Training, Simulation, and Education Conference (IITSEC)* (pp. 2924–2934). Orlando, FL: National Training and Simulation Association.
- Emond, B., & Vinson, N. G. (2017). Modelling Simple Ship Conning Tasks. In *15th Meeting of the International Conference on Cognitive Modelling* (pp. 1–2). Coventry, UK.
- Folsom-Kovarik, J. T., & Raybourn, E. M. (2016). Total Learning Architecture (TLA) Enables Next-generation Learning via Meta-adaptation. In *Interservice/Industry Training, Simulation, and Education Conference (IITSEC) 2016* (pp. 1–13). Orlando, FL: National Training and Simulation Association. Retrieved from <https://www.adlnet.gov/adl-assets/uploads/2016/12/Total-Learning-Architecture-TLA-Enables-Next-generation-Learning-via-Meta-adaptation.pdf>
- Goodwin, G. A., Murphy, J. S., & Medford, A. L. (2016). Support for Training Effectiveness Assessment and Data Interoperability. In *Interservice/Industry Training, Simulation, and Education Conference (IITSEC)* (pp. 1–11).
- Gupta, R. (2012). Risk Assessment Techniques and Survey Method for COTS Components. *International Journal of Software Engineering & Applications, 3*(1), 181–195. <http://doi.org/10.5121/ijsea.2012.3113>
- Hruska, M., Long, R., & Amburn, C. (2014). Human performance interoperability via xAPI: Current military outreach efforts. *Fall Simulation Interoperability Workshop, 2014 Fall SIW*, 207–216. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-84910145843&partnerID=tZOtx3y1>
- IMS Global Learning Consortium. (2015). Caliper Analytics. Retrieved August 10, 2017, from <http://www.imsglobal.org/activity/caliper>
- Poeppelman, T. R., Ayers, J., Hruska, M., Long, R., Amburn, C., & Bink, M. (2013). Interoperable Performance Assessment using the Experience API. *Interservice/Industry Training*, (13273).
- Simulation Interoperability Standards Organization. (2016a). Human Performance Markup Language (HPML) Development Group (PDG). Retrieved from <https://www.sisostds.org/StandardsActivities/DevelopmentGroups/HPMLPDG-HumanPerformanceMarkupLanguage.aspx>
- Simulation Interoperability Standards Organization. (2016b). Product Nomination for Human Performance Markup Language (SISO-PN-015-2016). Retrieved from https://www.sisostds.org/DigitalLibrary.aspx?Command=Core_Download&EntryId=43990
- Stacy, W., Ayers, J., Freeman, J., & Haimson, C. (2006). Representing human performance with human performance measurement language. In *Proceedings of the Fall 2006 Simulation Interoperability Workshop* (pp. 570–580). SISO - Simulation Interoperability Standards Organization. Retrieved from http://www.aptima.com/publications/2006_Stacy_Ayers_Freeman_Haimson.pdf%5Cnhttp://www.sisostds.org/StandardsActivities/StudyGroups/HPMLSGHumanPerformanceMarkupLanguage.aspx
- Stacy, W., & Freeman, J. (2016). Training objective packages: enhancing the effectiveness of experiential training. *Theoretical Issues in Ergonomics Science, 17*(2), 149–168. <http://doi.org/10.1080/1463922X.2015.1111459>
- Stacy, W., Merket, D., Freeman, J., Wiese, E., & Jackson, C. (2005). A Language for Rapidly Creating Performance Measures in Simulators. In *The Interservice/Industry Training, Simulation & Education Conference (IITSEC)*. Retrieved from <http://ntsa.metapress.com/index/3DFXHFCX6VTYFJPE.pdf>
- Walker, A., Tolland, M., & Stacy, W. (2015). Using a Human Performance Markup Language for simulator-based training. In *Simulation Interoperability Standards Organization (SISO) Simulation Interoperability Workshop (SIW) Meeting*. Orlando, FL: SISO - Simulation Interoperability Standards Organization.