

xAPI Data in Action | 17 August

Don't Just Give Me "All the Data" - Using xAPI for Training Effectiveness Evaluation

Jason Haag

Veracity Technology Consultants

iFEST
#ADLiFEST





Federal E-Learning Science & Technology

³**iFEST**

innovation • instruction • implementation

Don't Just Give Me "All the Data"

Using xAPI for Training Effectiveness Evaluation



Jason Haag, Co-Founder & Consultant

- Evaluation Strategy
- Measurement Strategy
- Aligning xAPI to Your Strategies
 - xAPI Profiles
 - Measures of Effectiveness (KM Levels 1-3)
 - Measures of Efficiency (TDRp) Level 0
- Resources

Start with the end in mind: Evaluation
“The End is the Beginning”



TEEP – OPNAV INST 1500.76D



DEPARTMENT OF THE NAVY
OFFICE OF THE CHIEF OF NAVAL OPERATIONS
2000 NAVY PENTAGON
WASHINGTON DC 20350-2000

OPNAVINST 1500.76D
N13M
29 Jul 2021

OPNAV INSTRUCTION 1500.76D

From: Chief of Naval Operations

Subj: NAVAL TRAINING SYSTEMS REQUIREMENTS, ACQUISITION AND
MANAGEMENT

Ref: (a) OPNAVINST 5420.117
(b) DoD Instruction 5000.80 of 30 December 2019
(c) OPNAVINST 1500.27G
(d) SECNAVINST 5000.2F
(e) CCSI 5123.01H
(f) OPNAVINST 5310.23A
(g) DOD Instruction 5000.75 of 2 February 2017
(h) OPNAVINST 1540.56B
(i) SECNAVINST 11011.47D
(j) OPNAVINST 3500.34G
(k) NAVSO P-1000 of November 2019
(l) SECNAVINST 5000.36A
(m) DoD Directive 5000.59 of 8 August 2007
(n) OPNAVINST 11010.20H
(o) OPNAVINST 4790.4F
(p) COMUSFLFORCOMINST/COMPACFLTINST 3501.3D
(q) OPNAVINST 4790.16B
(r) MIL-STD-46855a, Human Engineering Requirements for Military Systems,
Equipment and Facilities, 24 May 2011
(s) OPNAVINST 1000.16L
(t) DoD Instruction 5000.02T of 7 January 2015
(u) OPNAVINST 3500.23D

Encl: (1) Manpower and Training Requirements Planning
(2) Training Estimate and Navy Training System Plan Format
(3) Training Effectiveness Evaluation Plan
(4) Training Installation and Transfer Agreement
(5) Rating Continuum
(6) Glossary and Acronyms

OPNAVINST 1500.76D
29 Jul 2021

TRAINING EFFECTIVENESS EVALUATION PLAN

1. **Introduction.** The TEEP identifies and describes the plan to conduct a Training Effectiveness Evaluation (TEE) of the KPP or KSA (or both), critical tasks and associated standards related to the training delivered. The TEEP provides a strategy to collect and evaluate performance data of E/S/S, training systems, curricula and sailors.
2. **Purpose.** To document the strategy, schedule and measures of effectiveness as determined by the TYCOMs, NETC and other stakeholders.
3. **Process.** The PM is responsible for the TEEP in coordination with NETC and TYCOM. At a minimum, a TEEP will be developed in alignment with the MTRP or test and evaluation master plan (TEMP) (or both) and builder's trial cards resulting from changes to a Rating Continuum. Updates will be developed based on MTRP triggers and incorporated into the next scheduled TEE. A TEE will be conducted no earlier than six months after completion of the pilot training event and when the system and fleet occupational performance data is available to collect and evaluate. The results of the TEE may drive revisions to MTRP products (FEA, RDA, etc.)
4. **Procedure.** The actions identified in subparagraph 4a-c will be performed to create a TEEP and TEE.
 - a. The task analysis (system or rating) data is the authoritative source for the system and learning objectives linked to occupational standards providing the foundational starting point for planning.
 - b. Create an individual detailed evaluation plan for the TEEA, who will conduct the TEE. Identify the resources required to conduct the TEE. The TEEP requires TYCOM concurrence.
 - c. The TEEP will have the following sections: front matter, program identification, data collection strategy, evaluation strategy, TEEA selection and TEE reporting.
 - (1) Front Matter. Includes a cover page that identifies the training system or the training device, date prepared and activity or organization by code preparing the document and a table of contents identifying training system or training device name and number.
 - (2) Part 1 - Program Identification. Identify the host activity and location of the training site(s). List identification of TEEA(s) to send team members with lead activity also identified. List the number, rank, rate and rating of students to evaluate. List the number and type of

- Start with the end in mind – Evaluation
 - What are the reporting goals/requirements?
 - informs why you want to measure

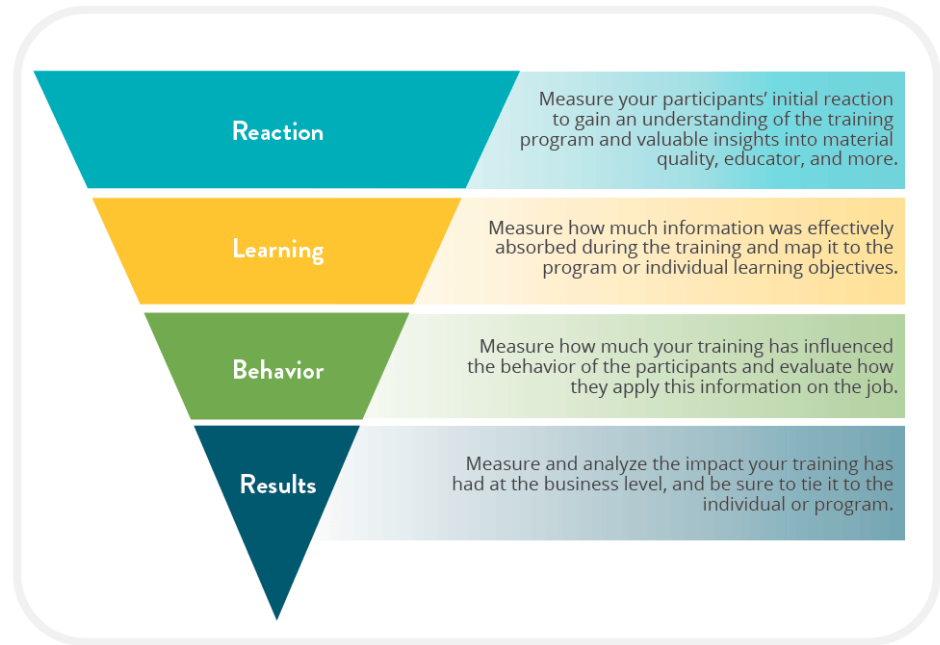


- Start with the end in mind – Evaluation
 - What are the reporting goals/requirements?
 - informs why you want to measure
 - Who is the org/client/customer and/or decision-makers?
 - they should help define the reporting goals/requirements



- Start with the end in mind – Evaluation
 - What are the reporting goals/requirements?
 - informs why you want to measure
 - Who is the org/client/customer and/or decision-makers?
 - they should help define the reporting goals/requirements
 - What model/framework can you use/tweak for your strategy?
 - Kirkpatrick, Phillips ROI, Talent Development Reporting Principles Framework (**TDRp**), etc.

THE KIRKPATRICK MODEL



Talent Development Reporting Principles (TDRp)

<https://www.centerfortalentreporting.org>

- Established to provide L&D and HR standards for reporting & analysis, similar to Generally Accepted Accounting Principles (GAAP) and International Financial Reporting Standards (IFRS)
- Helps you figure out why and reasons for measuring
- Focuses on not just Effectiveness, but also Efficiency and Outcome measures



Measurement Strategy (Infusing TDRp)

- Build the Evaluation Strategy
 - Why do you want to measure?



Measurement Strategy (Infusing TDRp)

- Build the Evaluation Strategy
 - Why do you want to measure?
- Reasons for Measuring:
 - Inform, Monitor, Evaluate & Analyze, Manage



Measurement Strategy (Infusing TDRp)

- Build the Evaluation Strategy
 - Why do you want to measure?
- Reasons for Measuring:
 - Inform, Monitor, Evaluate & Analyze, Manage
- Types of Measures:
 - Efficiency, Effectiveness, Outcome (Results)



- A profile is the “blueprint” for a successful xAPI implementation!
- Provides a common vocabulary, descriptive metadata, and the data collection requirements (what should be captured/stored and reported on).
- Reuse existing profiles when you can!
 - Courses (i.e., SCORM, cmi5)
 - Assessment
 - Serious Games
 - Audio/Video
 - E-books
 - Performance Support
 - Simulations
- ADL’s xAPI Profile Server: <https://profiles.adlnet.gov>
- Torrance Learning: <https://torrancelearning.com/xapi/profiles/tools/flattener>

EXPERIENCE xAPI



Measures of Effectiveness

Aligning xAPI with the Kirkpatrick Model Levels 1-3



Align Reporting Goals & Measures w/ xAPI

Reporting Goal	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	Assessment Technique	xAPI Data Collection Req.
What is the reporting goal/requirement identified/approved by the stakeholders?	What will be measured to provide data for the reporting goal?	What is the purpose of the measure?	What feedback will be provided from the measure? (think about the reasons - inform, monitor, analyze, manage).	What types of reports or analytics will be created?	What types of assessment instruments will be used?	What xAPI data should be collected to support the measure and reporting goal?



KM Level 1 - Reaction

Reporting Goal	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	Assessment Technique	xAPI Data Collection Req.
Instructor Satisfaction (4.1 – 5.0)	Favorable user satisfaction with the instructor	Measures how the users feel about the instructor and if the instructor is performing above the benchmark	<ul style="list-style-type: none">- Informs stakeholders how the instructor is performing- Helps stakeholders make informed decisions about operations and resourcing	- Instructor Satisfaction Report	Survey	Assessment Profile
Content Satisfaction (4.1 – 5.0)	Favorable user satisfaction with the content	Measures how the users feel about the content design, subject matter, modality, or resources, and if it is performing above the benchmark	<ul style="list-style-type: none">- Informs stakeholders how the content is performing- Helps stakeholders make decisions about making changes or improvements to the training content, curriculum, or other resources	- Learning Content Satisfaction Report	Survey	Assessment Profile
Learning Environment Satisfaction (4.1 – 5.0)	Favorable user satisfaction with the learning environment , resources, or systems	Measures how the users feel about the quality and appropriateness of the environment, ease of access, and if it is performing above the benchmark	<ul style="list-style-type: none">- Informs stakeholders how the environment is performing- Helps stakeholders make improvements to the learning environment	- Learning Environment Satisfaction Report	Survey	Assessment Profile



KM Level 2 - Learning

Reporting Goal	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	Assessment Technique	xAPI Data Collection Req.
Pretest Pass Rate	Number of users that passed a knowledge assessment pretest	Measures the users' existing knowledge before completing a learning activity	<ul style="list-style-type: none">- Informs stakeholders of users' existing knowledge of subject matter- Informs stakeholders of users' performance above or below the passing score threshold- Provides benchmark for later comparison against the posttest results	Pretest Pass/Fail Rate Report, Scorecard, or Dashboard	Pretest Knowledge Assessment	Assessment Profile
Practice Test Pass Rate	Number of users that passed or failed a knowledge assessment quiz or practice test	Measures the users' new knowledge acquired during a learning activity	<ul style="list-style-type: none">- Informs stakeholders of users' increased knowledge- Informs stakeholders of high and low performers, above or below the passing score threshold- Provides a metric for comparing against the pretest results	<ul style="list-style-type: none">- Posttest Pass/Fail Rate Report, Scorecard, or Dashboard- Pretest/Posttest Comparison Report, Scorecard, or Dashboard	Practice/Quiz Knowledge Assessment	Assessment Profile
Posttest Pass Rate	Number of users that passed or failed a knowledge assessment posttest	Measures the users' increase in knowledge after completing the learning activity	<ul style="list-style-type: none">- Informs stakeholders of users' increased knowledge- Informs stakeholders of high and low performers, above or below the passing score threshold- Provides a metric for comparing against the pretest results	<ul style="list-style-type: none">- Posttest Pass/Fail Rate Report, Scorecard, or Dashboard- Pretest/Posttest Comparison Report, Scorecard, or Dashboard	Posttest Knowledge Assessment	Assessment Profile



KM Level 2 – Learning (Cont'd)

Reporting Goal	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	Assessment Technique	xAPI Data Collection Req.
Refresher Pass Rate	Number of users that passed or failed a refresher knowledge assessment	Measures retainment and knowledge decay after completing the refresher learning activity	<ul style="list-style-type: none">- Informs stakeholders of users' knowledge retention or decrease in knowledge- Informs and monitors users' performance above or below the passing score threshold- Provides a metric for comparison against previous assessment results	<ul style="list-style-type: none">- Refresher Pass/Fail Rate Report, Scorecard, or Dashboard- Posttest/Refresher Comparison Report, Scorecard, or Dashboard	Refresher Knowledge Assessment	Assessment Profile
Attempts Made Before Passing	Number of attempts users made before they passed the knowledge assessment	Measures when a user's first successful attempt on an assessment meets or exceeds the threshold score	<ul style="list-style-type: none">- Informs stakeholders of a user's knowledge attainment speed-Informs stakeholders of potential improvements to the quality of the learning activity- Informs stakeholders to check for users that keep attempting the assessment until they pass (or exceed the limit)	<ul style="list-style-type: none">- First Attempt Pass Rate Report, Scorecard, or Dashboard- Time to Knowledge Attainment Report, Scorecard, or Dashboard	Knowledge Assessment	Assessment Profile
≥ 80% Combined Average Passing Score	Average (mean) score of all users that passed or failed a knowledge assessment	Measures the average of all users' scores to be represented as the standard mean score	<ul style="list-style-type: none">- Provides stakeholders with a benchmark to compare user's individual scores to the standard mean score- Provides stakeholders with a benchmark to compare the standard mean score to the threshold score	Average Score Report, Scorecard, or Dashboard	Knowledge Assessment	Assessment Profile



KM Level 2 – Learning (Cont'd)

Reporting Goal	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	Assessment Technique	xAPI Data Collection Req.
≥ 80% First Attempt Average Passing Score	Average (mean) score of all users that passed a knowledge assessment on the first attempt	Measures the average of all users' scores that passed on the first attempt to be represented as the standard first attempt passing mean score	<ul style="list-style-type: none">- Provides stakeholders with a benchmark metric to compare the standard mean score to the standard first attempt passing mean score- Provides stakeholders with a benchmark metric to compare the standard first attempt passing mean score to the threshold score	-First-Time Average Score Report, Scorecard, or Dashboard	Knowledge Assessment	Assessment Profile
Two Standard Deviations Above or Below the Average	Standard deviation (SD) from the average score of all users that passed a knowledge assessment	Measures how much variance there is from the standard mean score	<ul style="list-style-type: none">- Provides stakeholders with a metric to indicate a low SD (where the data points are close to the mean)- Provides stakeholders with an outlier metric to indicate a high SD (where the data are spread out over a large range of scores)	<ul style="list-style-type: none">-Low Standard Deviation Report, Scorecard, or Dashboard- High Standard Deviation Report, Scorecard, or Dashboard	Knowledge Assessment	Assessment Profile
Test Item Analysis	Number of correct and incorrect responses on a knowledge assessment	Measures validity of individual test questions in an assessment	<ul style="list-style-type: none">- Informs stakeholders of potential training effectiveness problems and revisions needed to the learning objectives, training activity, and assessment questions	-Test Item Analysis Report, Scorecard, or Dashboard	Knowledge Assessment	Assessment Profile



KM Level 2 – Learning (Cont'd)

Reporting Goal	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	Assessment Technique	xAPI Data Collection Req.
Self-Reported Assessment Score (4.1 – 5.0)	Opinions or beliefs (attitudes) of users about the how much they learned	Measures the user's self-assertion of their increase in knowledge or skills after completing the training activity	- Informs stakeholders of feedback about whether the users feel the training was effective for future on-the-job performance	-Learning Attitudes Report -Learning Attitudes Dashboard - Survey Responses Export (LRS) as CSV, JSON	Survey	Assessment Profile
Observational Assessment Score (4.1 – 5.0)	Instructor observation and verification on a performance assessment	Measures the instructor's assertion of the user's successful or unsuccessful performance on a classroom lab exercise, drill, or hands-on skill activity	- Provides stakeholders with additional evidence of the user's skills or performance of a task	-Observational Performance Assessment Scores Report, Scorecard, or Dashboard	Survey	Performance Assessment Profile
≥ 85% Scenario-Based Assessment Score	Number of problems recognized, correct actions taken, or correct sequence of actions by a user in a scenario-based simulation or assessment	Measures the users' decision-making or problem-solving skills after completing the training activity	- Informs stakeholders of users' increased decision-making skills or abilities - Informs stakeholders of high and low performers - Provides a metric for comparing against other results	-Decision-Making Performance Report, Scorecard, or Dashboard	Scenario-Based Assessment	- Assessment Profile - Performance Assessment Profile



KM Level 3 – Behavior

Reporting Goal	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	Assessment Technique	xAPI Data Collection Req.
Self-Reported Performance Assessment Score (4.1 – 5.0)	Confidence (attitudes) of users about their intent to apply/perform	Measures how the users intend to apply on-the-job what was previously learned via training	<ul style="list-style-type: none">- Informs stakeholders if the user perceives their confidence in performing- Helps stakeholders make decisions about making changes or improvements to the training- Provides a metric for comparing against Level 2 (Learning) for training effectiveness	- Intent to Apply Report	Survey	- Performance Assessment Profile
Supervisor Performance Assessment Score (4.1 – 5.0)	Supervisor observation and verification of performance	Measures the supervisor's assertion of the user's successful or unsuccessful performance on the job with actual equipment	<ul style="list-style-type: none">- Provides stakeholders with additional evidence of the user's skills or performance of a task	- Supervisor Performance Assessment Report	Survey	- Performance Assessment Profile

Measures of Efficiency

Aligning xAPI with TDRp (for Utilization, Reach, and Volume Metrics)



Level 0 - Measures of Efficiency (TDRo)

Modality	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	xAPI Data Collection Req.
-Instructor-Led Training (ILT) - Self-Paced Courses / ICW / SCORM	Number of total users	Measures the sum of total users for a training activity	This metric helps to monitor usage and when combined with other measures, it can inform future decisions about operations and resourcing.	- Total Users Report, Scorecard, or Dashboard - Total Users Over Time Dashboard	Actor Count
	Number of unique users	Measures the number of unique users for a training activity	This metric helps to monitor usage and when combined with other measures, it can inform future decisions about operations and resourcing.	- Unique Users Report, Scorecard, or Dashboard - Unique Users Over Time Dashboard	Actor Count
	Number of activities	Measures the total number of activities experienced for a training program	This metric helps to monitor usage and when combined with other measures, it can inform future decisions about operations and resourcing.	-Total Activities Report, Scorecard, or Dashboard	Object Count
	Number of activity types	Measures the total number of unique granular object types for a training activity	This metric helps to monitor usage and when combined with other measures, it can inform future decisions about operations and resourcing.	-Total Activity Types Report, Scorecard, or Dashboard	Activity Type Count



Level 0 - Measures of Efficiency (TDRp)

Modality	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	xAPI Data Collection Req.
-Instructor-Led Training (ILT) - Self-Paced Courses / ICW / SCORM	Number of views	Measures how many times a training activity was viewed	This metric helps to monitor current usage and can inform future decisions about the design.	- Total Views Report, Scorecard, or Dashboard - Total Views & Total Views Over Time Dashboard	Viewed Verb
	Duration / average time spent	Measures how much time was spent in a training activity	This metric can inform future decisions about the design, or it can be correlated with effectiveness measures (e.g., test performance, satisfying learning objectives).	Time Spent Report, Scorecard, or Dashboard Average Time Spent Report, Scorecard, or Dashboard	result.duration
	Number of played video segments	Measures how many times a training video was repeatedly played and how much of it was played	This metric can inform future decisions about the design, or it can be correlated with training effectiveness measures.	- Repeat Views Report, Scorecard, or Dashboard - Heat Map of Played Segments	Video Profile
	Number of completions	Measures the completion rate of a training activity (e.g., video), lesson, or course	This metric helps to monitor and compare against previous completion metrics / benchmarks.	- Completion Report, Scorecard, or Dashboard - Completion Over Time Dashboard	- SCORM Profile - Video Profile



Level 0 - Measures of Efficiency (TDRp)

Modality	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	xAPI Data Collection Req.
Performance Support	Number of unique and total users	Measures the sum of users accessing a performance support activity	This metric monitors usage and can inform future decisions about reach and access.	- Unique & Total Users Report, Scorecard, or Dashboard - Unique & Total Users Over Time Dashboard	Navy Common Reference Profile
	Number of page views and repeat views	Measures how many granular page views and repeat views	This metric monitors current usage and can inform future decisions about the design of the performance support solution.	Total Views & Repeat Views Report, Scorecard, or Dashboard	Navy Common Reference Profile
	Number of files opened	Measures how many times specific files were opened or downloaded	This metric monitors current usage and can inform future decisions about the design of the performance support solution.	Opened Files Report, Scorecard, or Dashboard	Navy Common Reference Profile
	Number of menu items accessed & repeat access	Measures how many times specific areas were accessed and repeatedly accessed	This metric monitors current usage and can inform future decisions about the design of the performance support solution.	Menu Item Access Report, Scorecard, or Dashboard	Navy Common Reference Profile



Level 0 - Measures of Efficiency (TDRp)

Modality	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	xAPI Data Collection Req.
Performance Support	Duration / average time spent	Measures how much time was spent in a performance support activity or resource	This metric monitors current usage and can inform future decisions about the design of the performance support solution.	Time Spent Report, Scorecard, or Dashboard	result.duration
	Number of played video segments	Measures how many times a training video was repeatedly played and how much of it was played	This metric monitors current usage and can inform future decisions about the design of the performance support.	- Total Views & Repeat Views Report, Scorecard, or Dashboard - Heat Map of Played Segments	Video Profile
	Number of video completions	Measures how many times a performance support video was completed	This metric informs future decisions about the design of the performance support solution and video resources provided.	Most Completed Videos Report, Scorecard, or Dashboard	Video Profile
	Number of liked/unliked resources	Measures how many times a performance support resource was liked and disliked	This metric informs future decisions about the design of the performance support solution and resources provided.	Most Liked & Unliked Report, Scorecard, or Dashboard	Navy Common Reference Profile



Level 0 - Measures of Efficiency (TDRp)

Modality	KPI / Measure	Purpose	Feedback Loop(s)	Reports/ Analytics	xAPI Data Collection Req.
Performance Support	Number of steps viewed and repeat steps viewed	Measures how many times specific steps in a procedure were viewed and repeatedly viewed	This metric monitors current usage and can inform future decisions about the design of the performance support solution.	Total Steps & Repeat Steps Report, Scorecard, or Dashboard	Performance Support Profile
	Number of procedures viewed and repeat procedures viewed	Measures how many times specific procedures were viewed and repeatedly viewed	This metric monitors current usage and can inform future decisions about the design of the performance support solution.	Total Procedures & Repeat Procedures Report, Scorecard, or Dashboard	Performance Support Profile
	Number of checklist completions and incompletions	Measures how many times a checklist was completed vs incomplete	This metric monitors current usage and can inform future decisions about the design of the performance support solution.	Checklist Completions Report, Scorecard, or Dashboard	Performance Support Profile
	Number of unique search terms and search results selected	Measures how many unique terms were searched for, and which search results were selected the most	This metric monitors current usage and can inform future decisions about the design of the performance support solution.	Search Terms Report, Scorecard, or Dashboard	Performance Support Profile

Summary (Using KM + TDRp + xAPI)

1. Define Reporting Purposes/Goals
 - Decision-makers identify the goals
 - Informed by KM Level 4 (results)
2. What model/framework can you use/tweak for your strategy?
 - Kirkpatrick, Phillips ROI, Talent Development Reporting Principles Framework (TDRp), etc.
3. Establish measures for the KM Levels
4. Align measures to xAPI Profiles or New xAPI Data Collection Requirements
5. Develop/Document the Evaluation Plan (e.g., TEEP, frequency) for the Training Product(s)



- Hall, Mattox, Parskey, *Learning Analytics: Using Talent Data to Improve Business Outcomes*
 - <https://amazon.com/Learning-Analytics-Improve-Business-Outcomes/dp/1789663008>
- Parskey and Vance, *Measurement Demystified: Creating Your L&D Measurement, Analytics, and Reporting Strategy*
 - <https://www.amazon.com/Measurement-Demystified-Creating-Analytics-Reporting/dp/1950496899>
- *Talent Development Reporting Principles (TDRp) Framework*,
 - <https://www.centerfortalentreporting.org>



Questions?

jason.haag@veracity.it

<https://veracity.it>

<https://lrs.io>

