

Designing High Quality Rubrics

A LIVETEXT™ Sponsored Presentation

At 2015 NYSATE/NYACTE Annual Fall Conference

Dr. Lance Tomei

Educational Consultant

Retired Director for Assessment, Accreditation, and Data Management, University of Central Florida, College of Education and Human Performance

Acknowledgement and Disclaimer

- Sincere thanks to LIVETEXT™!
- The content of this presentation reflects my **personal perspective** on the importance of designing high quality rubrics to ensure that resulting candidate performance data can be used effectively to improve candidate learning and program quality and thus meet the heightened expectations of CAEP.

Overview

- Reasons for the current focus on rubric quality
- Value-added of high quality rubrics
- Attributes of high quality rubrics
- An action plan

CAEP Standards: Underlying Principles

“CAEP Standards and their Components flow from two principles:

- **Solid evidence** that the provider’s graduates are competent and caring educators, and
- There must be **solid evidence** that the provider’s educator staff have the capacity to create a culture of evidence and use it to maintain and enhance the quality of the professional programs they offer.”

Introduction to CAEP Standards available at caepnet.org/standards/introduction

CAEP Standard 5, Component 5.2

“The provider’s quality assurance system relies on **relevant, verifiable, representative, cumulative and actionable** measures, and produces **empirical evidence that interpretations of data are valid and consistent** [*emphasis added*].”

Dr. Peter Ewell

Vice President, National Center for Higher Education Management Systems

Dr. Ewell has written numerous publications about the quality of evidence used to demonstrate student learning including papers for the Council for Higher Education Accreditation (CHEA), the National Institute for Learning Outcomes Assessment (NILOA), and the Council for the Accreditation of Educator Preparation (CAEP). In his article, “Principles for Measures Used in the CAEP Accreditation Process,” he suggests that all of the following qualities of evidence should be present:

1. Validity and Reliability
2. Relevance
3. Verifiability
4. Representativeness
5. Cumulativity
6. Fairness
7. Stakeholder Interest
8. Benchmarks
9. Vulnerability to Manipulation
10. Actionability

Article available at: <http://caepnet.org/standards/commission-on-standards>

Principles for Measures Used in the CAEP Accreditation Process

(Peter Ewell, May 29, 2013)

1. **Validity and Reliability** – relate to the fact that “All measures are in some way flawed and contain an error term that may be known or unknown.”
2. **Relevance** – “measures...ought to be demonstrably related to a question of importance that is being investigated.” (Why are you using this measure?)
3. **Verifiability** – “subject to independent verification...implies reliability...[plus] transparency and full documentation”
4. **Representativeness** – “sample is representative of the overall population”
5. **Cumulativeness** – “Measures gain credibility as additional sources or methods for generating them are employed . . . the entire set of measures used under a given Standard should be mutually reinforcing.”

Principles for Measures Used in the CAEP Accreditation Process

(Peter Ewell, May 29, 2013)

6. **Fairness** – “Measures should be free of bias and be able to be justly applied by any potential user or observer.”
7. **Stakeholder Interest** – “A sound set of measures should respect a range of client perspectives including the program, the student, the employer, and the state or jurisdiction.”
8. **Benchmarks** – “Without clear standards of comparison, the interpretation of any measure is subject to considerable doubt.”
9. **Vulnerability to Manipulation** – “All measures are to some extent vulnerable to manipulation. This is one reason to insist upon triangulation and mutual reinforcement across the measures used under each Standard.”
10. **Actionability** – “Good measures . . . should provide programs with specific guidance for action and improvement.”

Where Do We Stand?

“Many of the measures used to assess the adequacy of teacher preparation programs such as licensure examination scores meet these rigorous standards **but many of the more qualitative measures proposed do not**. Even the most rigorous measures, moreover, may not embrace the entire range of validities—construct, concurrent, and predictive.” (Ewell, 2013)

Optional CAEP Review of Assessment Instruments

CAEP allows the early submission of all key assessment instruments (rubrics, surveys, etc.) used by an Educator Preparation Provider (EPP) to generate data provided as evidence in support of CAEP accreditation. CAEP will evaluate these instruments and provide feedback to the EPP well prior to the formal accreditation review.

NOTE: CAEP has a draft document in development that includes rubrics they plan to use in their review of assessment instruments.

A Reality Check Regarding Current Rubrics: Commonly Encountered Weaknesses

- Using overly broad criteria
- Using double- or multiple-barreled criteria
- Using overlapping performance descriptors
- Failing to include all possible performance outcomes
- Using double-barreled descriptors that derail actionability
- Using subjective terms, performance level labels (or surrogates), or inconsequential terms to differentiate performance levels
- Failing to maintain the integrity of target learning outcomes: a common result of having multiple levels of “mastery”

Overly Broad Criterion

Criterion	Unsatisfactory	Developing	Proficient	Distinguished
Assessment	No evidence of review of assessment data. Inadequate modification of instruction. Instruction does not provide evidence of assessment strategies.	Instruction provides evidence of alternative assessment strategies. Some instructional goals are assessed. Some evidence of review of assessment data.	Alternative assessment strategies are indicated (in plans). Lessons provide evidence of instructional modification based on learners' needs. Candidate reviews assessment data to inform instruction.	Candidate selects and uses assessment data from a variety of sources. Consistently uses alternative and traditional assessment strategies. Candidate communicates with learners about their progress.

Double-barreled Criterion & Double-barreled Descriptor

Criterion	Unsatisfactory	Developing	Proficient
Alignment to Applicable State P-12 Standards and Identification of Appropriate Instructional Materials	Lesson plan does not reference P-12 standards or instructional materials.	Lesson plan references applicable P-12 standards OR appropriate instructional materials, but not both.	Lesson plan references applicable P-12 standards AND identifies appropriate instructional materials

Overlapping Performance Levels

Criterion	Unsatisfactory	Developing	Proficient	Distinguished
Communicating Learning Activity Instructions to Students	Makes two or more errors when describing learning activity instructions to students	Makes no more than two errors when describing learning activity instructions to students	Makes no more than one error when describing learning activity instructions to students	Provides complete, accurate learning activity instructions to students

Possible Gap in Performance Levels

Criterion	Unsatisfactory	Developing	Proficient	Distinguished
Instructional Materials	Lesson plan does not reference any instructional materials	Instructional materials are missing for one or two parts of the lesson	Instructional materials for all parts of the lesson are listed and directly relate to the learning objectives.	Instructional materials for all parts of the lesson are listed, directly relate to the learning objectives, and are developmentally appropriate.

Use of Subjective Terms

Criterion	Unsatisfactory	Developing	Proficient	Distinguished
Knowledge of Laboratory Safety Policies	Candidate shows a weak degree of understanding of laboratory safety policies	Candidate shows a relatively weak degree of understanding of laboratory safety policies	Candidate shows a moderate degree of understanding of laboratory safety policies	Candidate shows a high degree of understanding of laboratory safety policies

Use of Performance Level Labels

Criteria	Unacceptable	Acceptable	Target
Analyze Assessment Data	Fails to analyze and apply data from multiple assessments and measures to diagnose students' learning needs, inform instruction based on those needs, and drive the learning process in a manner that documents acceptable performance.	Analyzes and applies data from multiple assessments and measures to diagnose students' learning needs, informs instruction based on those needs, and drives the learning process in a manner that documents acceptable performance.	Analyzes and applies data from multiple assessments and measures to diagnose students' learning needs, informs instruction based on those needs, and drives the learning process in a manner that documents targeted performance.

Use of Surrogates for Performance Levels

Criterion	Unsatisfactory	Developing	Proficient	Distinguished
Quality of Writing	Poorly written	Satisfactorily written	Well written	Very well written

Use of Inconsequential Terms

Criteria	Unacceptable	Acceptable	Target
Alignment of Assessment to Learning Outcome(s)	The content of the test is not appropriate for this learning activity and is not described in an accurate manner.	The content of the test is appropriate for this learning activity and is described in an accurate manner.	The content of the test is appropriate for this learning activity and is clearly described in an accurate manner.

Failure to Maintain Integrity of Target Learning Outcomes

Criterion	Unsatisfactory	Developing	Proficient	Distinguished
Alignment to Applicable State P-12 Standards	No reference to applicable state P-12 standards	Referenced state P-12 standards are not aligned with the lesson objectives and are not age-appropriate	Referenced state P-12 standards are age-appropriate but are not aligned to the learning objectives.	Referenced state P-12 standards are age-appropriate and are aligned to the learning objectives.

Criterion	Unsatisfactory	Developing	Proficient	Distinguished
Instructional Materials	Lesson plan does not reference any instructional materials	Instructional materials are missing for one or two parts of the lesson	Instructional materials for all parts of the lesson are listed and relate directly to the learning objectives.	Instructional materials for all parts of the lesson are listed, relate directly to the learning objectives, and are developmentally appropriate.

Value-Added of High Quality Rubrics

For Candidates, Well-designed Rubrics Can:

- Serve as an effective learning scaffold by clarifying formative and summative learning objectives (i.e., clearly describing expected performance at key formative stages and at program completion)
- Identify the critical indicators aligned to applicable standards/ competencies (=construct & content validity) for each target learning outcome
- Facilitate self- and peer-evaluations

Value-Added of High Quality Rubrics

For Faculty, Well-designed Rubrics Can:

- Improve assessment of candidates' performance by:
 - Providing a consistent framework for key assessments
 - Ensuring the consistent use of a set of critical indicators (i.e., the rubric criteria) for each competency
 - Establishing clear/concrete performance descriptors for each assessed criterion at each performance level
- Help ensure strong articulation of formative and summative assessments
- Improve validity and inter- and intra-rater reliability of assessment data
- Produce actionable candidate- and program-level data

Common 4-Level Rubric Template

Criteria	Unsatisfactory (0 pts)	Developing (1 pt)	Proficient (2 Pts)	Exemplary (3 pts)
Criterion #1 [Standard(s)]				
Criterion #2 [Standard(s)]				
Criterion #3 [Standard(s)]				
Criterion #4 [Standard(s)]				

Better Option #1: 4-Level Rubric

(Two formative assessments)

Criteria	Unsatisfactory (0 pts)	Developing 1 (1 pt)	Developing 2 (2 Pts)	Mastery (3 pts)
Criterion #1 [Standard(s)]				
Criterion #2 [Standard(s)]				
Criterion #3 [Standard(s)]				
Criterion #4 [Standard(s)]				

AAC&U VALUE Rubric – Information Literacy

Available online at http://www.aacu.org/value/rubrics/index_p.cfm

INFORMATION LITERACY VALUE RUBRIC

for more information, please contact valu@aacu.org



Definition

The ability to know when there is a need for information, to be able to identify, locate, evaluate, and effectively and responsibly use and share that information for the problem at hand. - The National Forum on Information Literacy

Evaluators are encouraged to assign a zero to any work sample or collection of work that does not meet benchmark (all one) level performance.

	Capstone 4	Milestones		Benchmark 1
		3	2	
Determine the Extent of Information Needed	Effectively defines the scope of the research question or thesis. Effectively determines key concepts. Types of information (sources) selected directly relate to concepts or answer research question.	Defines the scope of the research question or thesis completely. Can determine key concepts. Types of information (sources) selected relate to concepts or answer research question.	Defines the scope of the research question or thesis incompletely (parts are missing, remains too broad or too narrow, etc.). Can determine key concepts. Types of information (sources) selected partially relate to concepts or answer research question.	Has difficulty defining the scope of the research question or thesis. Has difficulty determining key concepts. Types of information (sources) selected do not relate to concepts or answer research question.
Access the Needed Information	Accesses information using effective, well-designed search strategies and most appropriate information sources.	Accesses information using variety of search strategies and some relevant information sources. Demonstrates ability to refine search.	Accesses information using simple search strategies, retrieves information from limited and similar sources.	Accesses information randomly, retrieves information that lacks relevance and quality.
Evaluate Information and its Sources Critically*	Chooses a variety of information sources appropriate to the scope and discipline of the research question. Selects sources after considering the importance (to the researched topic) of the multiple criteria used (such as relevance to the research question, currency, authority, audience, and bias or point of view).	Chooses a variety of information sources appropriate to the scope and discipline of the research question. Selects sources using multiple criteria (such as relevance to the research question, currency, and authority).	Chooses a variety of information sources. Selects sources using basic criteria (such as relevance to the research question and currency).	Chooses a few information sources. Selects sources using limited criteria (such as relevance to the research question).
Use Information Effectively to Accomplish a Specific Purpose	Communicates, organizes and synthesizes information from sources to fully achieve a specific purpose, with clarity and depth	Communicates, organizes and synthesizes information from sources. Intended purpose is achieved.	Communicates and organizes information from sources. The information is not yet synthesized, so the intended purpose is not fully achieved.	Communicates information from sources. The information is fragmented and/or used inappropriately (misquoted, taken out of context, or incorrectly paraphrased, etc.), so the intended purpose is not achieved.
Access and Use Information Ethically and Legally	Students use correctly all of the following information use strategies (use of citations and references; choice of paraphrasing, summary, or quoting; using information in ways that are true to original context; distinguishing between common knowledge and ideas requiring attribution) and demonstrate a full understanding of the ethical and legal restrictions on the use of published, confidential, and/or proprietary information.	Students use correctly three of the following information use strategies (use of citations and references; choice of paraphrasing, summary, or quoting; using information in ways that are true to original context; distinguishing between common knowledge and ideas requiring attribution) and demonstrates a full understanding of the ethical and legal restrictions on the use of published, confidential, and/or proprietary information.	Students use correctly two of the following information use strategies (use of citations and references; choice of paraphrasing, summary, or quoting; using information in ways that are true to original context; distinguishing between common knowledge and ideas requiring attribution) and demonstrates a full understanding of the ethical and legal restrictions on the use of published, confidential, and/or proprietary information.	Students use correctly one of the following information use strategies (use of citations and references; choice of paraphrasing, summary, or quoting; using information in ways that are true to original context; distinguishing between common knowledge and ideas requiring attribution) and demonstrates a full understanding of the ethical and legal restrictions on the use of published, confidential, and/or proprietary information.

*Corrected Dimension 3: Evaluate Information and its Sources Critically in July 2013

4-Level Rubric Template

One Possible Conceptual Framework

Criteria	Unsatisfactory	Remembering, Understanding	Applying	Analyzing, Evaluating, Creating
		Demonstration of content and pedagogical knowledge	Informed practice	Reflective and impactful practice
Criterion 1				
Criterion 2				
Criterion 3				
Criterion 4				

Attributes of an Effective Rubric

1. Rubric and the assessed activity or artifact are well-articulated.

Attributes of an Effective Rubric (cont.)

2. Rubric has construct validity (i.e., it measures the right competency/ies) and content validity (rubric criteria represent all *critical* indicators for the competency/ies to be assessed).

How do you ensure this?

Attributes of an Effective Rubric (cont.)

3. Each criterion assesses an individual construct
 - No overly broad criteria
 - No double- or multiple-barreled criteria

Attributes of an Effective Rubric (cont.)

4. To enhance reliability, performance descriptors should:

- Provide clear/concrete distinctions between performance levels (there is no overlap between performance levels)
- Collectively address all possible performance levels (there is no gap between performance levels)
- Eliminate (or carefully construct the inclusion of) double/multiple-barrel narratives

Attributes of an Effective Rubric (cont.)

5. Contains no unnecessary performance levels.

Common problems encountered when **multiple levels of mastery** are present include:

- Use of subjective terms to differentiate performance levels
- Use of performance level labels or surrogates
- Use of inconsequential terms to differentiate performance levels
- **Worst case scenario: failure to maintain the integrity of target learning outcomes**

Attributes of an Effective Rubric

6. Resulting data are actionable

- To remediate individual candidates
- To help identify opportunities for program quality improvement

Based on the first four attributes, the following meta-rubric has been developed for use in evaluating the efficacy of other rubrics...

“Meta-rubric” to Evaluate Rubric Quality

Criteria	Unsatisfactory	Developing	Mastery
Rubric Alignment to Assignment.	The rubric includes multiple criteria that are not explicitly or implicitly reflected in the assignment directions for the learning activity to be assessed.	The rubric includes one criterion that is not explicitly or implicitly reflected in the assignment directions for the learning activity to be assessed.	The rubric criteria accurately match the performance criteria reflected in the assignment directions for the learning activity to be assessed.
Comprehensiveness of Criteria	More than one critical indicator for the competency or standard being assessed is not reflected in the rubric.	One critical indicator for the competency or standard being assessed is not reflected in the rubric.	All critical indicators for the competency or standard being assessed are reflected in the rubric.
Integrity of Criteria	More than one criterion contains multiple, independent constructs (similar to “double-barreled survey question”).	One criterion contains multiple, independent constructs. All other criteria each consist of a single construct.	Each criterion consists of a single construct.
Quality of Performance Descriptors	Performance descriptors are not distinct (i.e., mutually exclusive) AND collectively do not include all possible learning outcomes.	Performance descriptors are not distinct (i.e., mutually exclusive) OR collectively do not include all possible learning outcomes.	Performance descriptors are distinct (mutually exclusive) AND collectively include all possible learning outcomes.

Meta-Rubric Alignment to CAEP Expectations

Meta-rubric Criteria	CAEP Indicators of Rubric Quality per February 10, 2015 Draft Document "Rubrics for Evaluation of EPP Instruments Uses as Accreditation Evidence"
Rubric Alignment to Assignment	Instrument development is integrated with preparation curriculum
	The respondents for the instrument are given a description of its purpose
	Instructions provided to respondents about what they are expected to do are informative and unambiguous
	The purpose of the instrument and its use in candidate monitoring or decisions on progress are specified
Comprehensiveness of Criteria	The CAEP, InTASC, or state standards that the instrument will inform are explicit
	Instrument development engaged relevant preparation provider and clinical faculty
	The assessment items, or the assignment tasks: are consistent with the content of the standards being informed
	The assessment items, or the assignment tasks: represent the complexity or cognitive demands found in the standards
	The assessment items, or the assignment tasks: reflect the degree of difficulty or level of effort described in the standards
	Alignment criteria are consistently demonstrated (50% to 75%)
	Assessments and assignments include items congruent with standard/components that require higher level of intellectual behavior . . .
	Reviewer protocols contain evaluation categories clearly aligned with CAEP, InTASC, and/or state standards
Integrity of Criteria	Most evaluation categories (80% of the total score) require observers to judge condequential attributes of candidate proficiencies in the standards
	Feedback provided to candidates is actionable
	Performance attributes are defined in actionable, performance based, or observable behavior terms If a less actionable term is used such as "engaged," criteria are provided to define the use of the term in the context of the item
Quality of Performance Descriptors	The basis for judgment (criterion for success, or what is "good enough" is made explicit for respondents
	Evaluation categories unambiguously describe the proficiencies to be evaluated
	Levels are qualitatively defined using specific criteria aligned with key attributes identified in the item
	Levels represent a developmental sequence from level to level. By qualitatively defining performance at each level, candidates are provided with descriptive feedback on their performance and consistency across raters is increased.
	The basis for judging candidate work is well defined
	The point or points when the instrument is administered during the preparation program are explicit
	The EPP provides a detailed description of the instrument's development
	Multiple raters or scorers are trained and used

Build a Solid Arch!

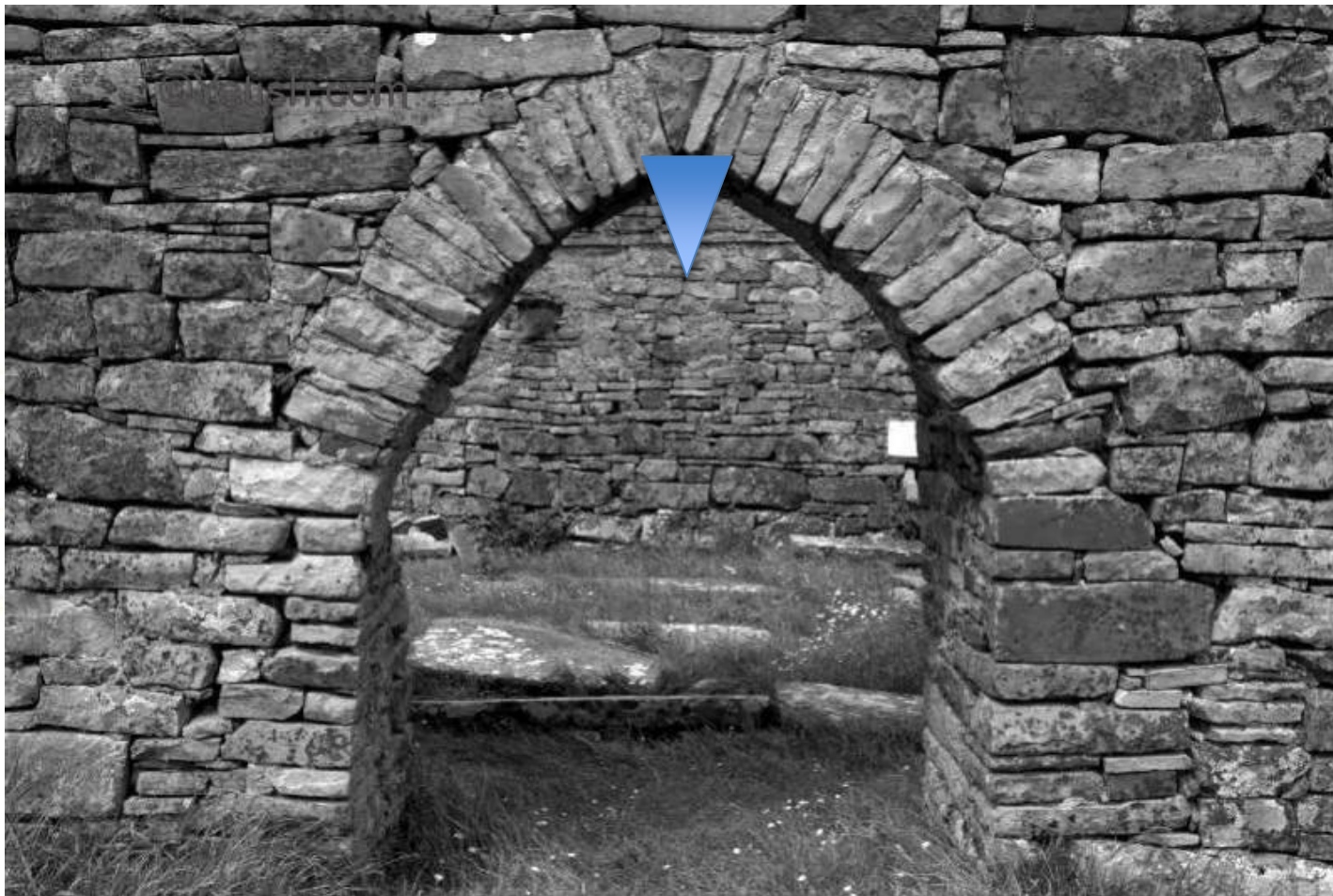


Image from <http://www.bing.com/image>

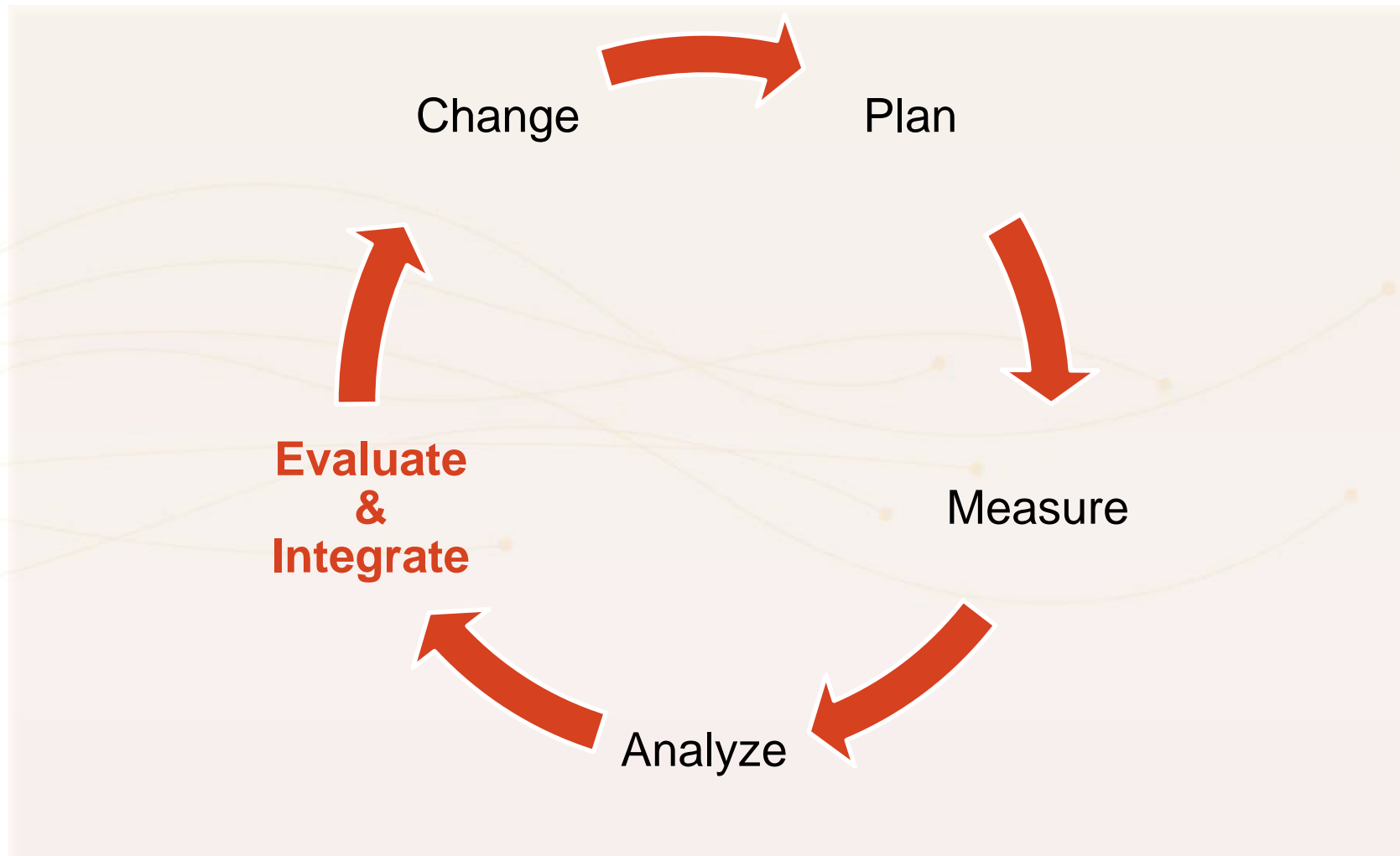
An Action Plan (Competency-based Approach)

1. Ensure that faculty are trained on rubric design.
2. Establish or review program-level target student learning outcomes.
3. For all target outcome performance competencies, identify the observable indicators that you consider the “must see” indicators (aka “critical indicators”) to establish solid evidence of mastery. Consider both effectiveness and efficiency when completing this task!
4. Determine the number of times you want to assess these competencies, both formatively and summatively. Your rubric should include this number of performance levels, plus one.
5. Critical indicators for each competency become the assessed elements in your rubric(s).

An Action Plan (Continued)

5. First, populate the performance descriptors for all critical indicators at the target level. This will ensure that you maintain the integrity of your target learning outcomes. **If you include multiple levels of mastery, populate the lowest level of mastery first.**
6. Populate the performance descriptors for formative levels next, from highest to lowest.
7. Define what would be considered “Unacceptable” or “Unsatisfactory” at the first formative assessment, and include those descriptors.
8. Populate higher levels of mastery (if used) with descriptors that are differentiated from the target learning outcomes in a clear, concrete manner.
9. Train faculty on the use of these rubrics and conduct reliability testing.

The Continuous Quality Improvement Cycle



LIVETEXT™ Visitor Pass

- Go to www.livetext.com
- Click on “Visitor Pass”
- Enter Visitor Pass code **“856EB982”**
- Click on Visitor Pass Entry, then click on “Designing High Quality Rubrics”
- You will have access to
 - This PowerPoint presentation
 - My “meta-rubric”
 - CAEP rubrics for evaluating assessment instruments (DRAFT)
 - Crosswalk of my meta-rubric with indicators in the CAEP rubrics
 - A link to LIVETEXT archived Webinars
 - Links to AAC&U VALUE rubrics and other useful rubric design resources
 - Principles for Measures Used in the CAEP Accreditation Process (Peter Ewell, May 29, 2013)
 - CAEP Evidence Guide January 2015
 - CAEP Accreditation Manual DRAFT dated February 2, 2015
 - Links to the latest versions of CAEP standards for initial and advanced programs
 - InTASC Model Core Teaching Standards 2011



Questions/Comments?