

# Development and Analysis of EPP-Created Assessments

---

*Adapted from CAEP Workshop on April 18, 2017 in Atlanta, GA*

# Goal and Objectives

**Goal:** The goal of the upcoming program coordinator meetings is to provide program coordinators with an opportunity to review the quality of EPP- or program-created assessments and make improvements.

## **Objectives:**

- Describe the criteria of the CAEP Evaluation Framework as it applies to EPP or program assessment(s).
- Apply the criteria of the CAEP Evaluation Framework and a description of common weaknesses to a self-critique of current assessment(s).
- Revise at least one assessment to meet the CAEP sufficient level.
- Generate a list of next steps for further revisions and/or EPP discussion and planning.



# Topics Covered

- Administration and purpose
- Content of assessment
- Scoring

# Overview of the CAEP Evaluation Framework (CEF)

- The CEF applies to any EPP-created assessment: rubrics, surveys, etc.
- The CEF does not apply to any proprietary assessment: edTPA, licensure exams, Danielson framework, etc.
- CAEP no longer provides early instrument review; however, strongly recommends using the CEF with internal or external experts conducting the review.
- The CEF is one tool used by site visitors at the point of the off-site review and the Formative Feedback Report to evaluate the quality of EPP-created assessments and data trustworthiness (validity and reliability).

# Administration and Purpose

<b><u>BELOW</u> Sufficient</b>	<b><u>SUFFICIENT</u> Level</b>	<b><u>ABOVE</u> Sufficient</b>
Use or purpose are ambiguous or vague.	The point or points when the assessment is administered during the preparation program are explicit.	
There is limited or no basis for reviewers to know what information is given to candidates.	The purpose of the assessment and its use in candidate monitoring or decisions on progression are specified and appropriate.	<ul style="list-style-type: none"> <li>a. The purpose of the assessments and its use in candidate monitoring or decisions are <u>consequential</u>.</li> <li>b. Candidate progression is monitored and information is used for <u>mentoring</u>.</li> </ul>
Instructions given to candidates are incomplete or misleading.	Instructions provided to candidates (or respondents to surveys) about what they are expected to do are informative and unambiguous.	
The criterion for success is not provided or is not clear.	The basis for judgment (criterion for success, or what is “good enough”) is made explicit for candidates (or respondents to surveys).	Candidates are informed <u>how the instrument results are used</u> in reaching conclusions about their <u>status</u> and/or <u>progression</u> .
	Evaluation categories or assessment tasks are aligned with CAEP, InTASC, national/professional and state standards.	



## Administration and Purpose Example 1: Statement at Below Sufficient Level

The observation rubric is introduced to candidates during Summer 1. It is used here for the instructors and the candidates to assess candidates' teaching in SSI. Not all of the standards on this rubric are appropriate in assessing the SSI teaching, so only certain standards are used. Candidates understand that the rubric assesses their teaching skills.

## Administration and Purpose Example 2: Statement at Sufficient Level

- The purpose of the rubric is to assess candidates developing skills in teaching, content knowledge as displayed during teaching, and dispositions as displayed in the learning environment. InTASC and NSTA standards guided the development of the rubric.
- The rubric is primarily used by the program faculty who teach academic courses and supervise candidates in their school placements twice each month during their ten month residency (September through June).
- Candidates are expected to achieve at the “basic” level of performance by mid-year (January) and at the “proficient” level by end-of-year (June).
- Observation evaluations from September through December, and again from January through May, are aggregated and used at the January and June Academic Milestone meetings when candidates’ standing in the program is evaluated and discussed.



# Tips for Administration and Purpose

- Consider the following:
  - Provide an overview statement that outlines:
    - When, how often and by whom the assessment is administered.
    - Why conduct this assessment? What is captured in terms of candidates' Knowledge, Skills, and Dispositions (KSD)?
    - What is the expected level of performance? At what points in the program?
    - How are the results of the assessment used?
    - What standards are used to guide content?
  - Include the candidate-facing instructions underlying the assignment/assessment.



# Task Workshop

\*Refer to Worksheet #1: Administration and Purpose

## **Tasks:**

- Draft a context statement using the assessment on Handout #1.
- Answer the questions on Worksheet #1.

# Rubric Terminology

Rubric Levels

Criteria/Indicators	Unacceptable	Acceptable	Proficient
Indicators	Proficiency-Level Descriptors		



# Content of Assessment

<b><u>BELOW</u> Sufficient</b>	<b><u>SUFFICIENT</u> Level</b>	<b><u>ABOVE</u> Sufficient</b>
Indicator alignment with CAEP, InTASC national/professional or state standards is incomplete, absent or only vaguely related to the content of the standards being evaluated	Indicators assess explicitly identified aspects of CAEP, InTASC, national/professional and state standards.	
Indicators fail to reflect the degree of difficulty described in the standard.	Indicators reflect the degree of difficulty or level of effort described in the standards.	
Indicators not described, are ambiguous, or include only headings.	Indicators unambiguously describe the proficiencies to be evaluated.	
Higher level functioning, as represented in the standards, is not apparent in the indicators.	When the standards being informed address higher level functioning, the indicators require higher levels of intellectual behavior (e.g., create evaluate, analyze & apply).	
Many indicators (more than 20% of the total score) require judgment of candidate proficiencies that are of limited importance in CAEP, InTASC, national/professional, and/or state standards.	Most indicators (at least those comprising 80% of the total score) require observers to judge consequential attributes of candidate proficiencies in the standards.	Almost all indicators (95% or more of the total score) require observers to judge consequential attributes of candidate proficiencies in the standards.

# Content of Assessment Example 1: Indicators at Below Sufficient Level

Indicators	Unacceptable	Acceptable	Proficient
Learning objectives			
Instructional strategies			
Assessment			



# Content of Assessment Example 2: Indicators at Sufficient Level

Indicators	Unacceptable	Acceptable	Proficient
1. <u>Develops</u> learning objectives which are appropriate for the subject and grade level that are connected to the standards. CAEP 1.1; InTASC 7			
2. <u>Plans</u> appropriate and logically sequenced instructional strategies. CAEP 1.1; InTASC 7			
3. <u>Plans</u> for differences in individual needs, abilities, and interests. CAEP 1.4: <b>InTASC 1</b>			

# Tips for Content of Assessment

Consider the following:

- Tag standards on the framework itself.
- Address InTASC categories by InTASC standard, not elements.

Beware the following:

- Double-dipped standards (InTASC 3 and 4).
- Headings as placeholders for indicators.



# Task Workshop

\*Worksheet #2: Content of Assessment: Indicators

## Task:

- Select at least two indicators from Handout #1 that do not currently meet the sufficient level.
- Rewrite the indicators by following the guidelines on Worksheet #2.
- Time permitting, evaluate the other indicators and make a list of needed revisions.

# Scoring

<b><u>BELOW</u> Sufficient</b>	<b><u>SUFFICIENT</u> Level</b>	<b><u>ABOVE</u> Sufficient</b>
Rating scales are used instead of rubrics.	The basis for judging candidate performance is well-defined.	
Proficiency Level Descriptors (PLDs) do not align with indicators.	Each Proficiency Level Descriptor (PLD) is qualitatively defined by specific criteria aligned with indicators.	Higher level actions from Bloom's or other taxonomies are used in PLDs such as "analyzes" or "evaluates".
PLDs do not represent developmental progressions.	PLDs represent a developmental sequence from level to level (to provide raters with explicit guidelines for evaluating candidate performance and for providing candidates with explicit feedback on their performance).	
PLDs provide limited or no feedback to candidates specific to their performance.	Feedback provided to candidates is actionable –it is directly related to the preparation program and can be used for program improvement as well as for feedback to the candidate.	
PLDS are vague or not defined, and may just repeat the language from the standards.	PLD attributes are defined in actionable, performance-based, or observable behavior terms.	



# Scoring Example 1:

## Proficiency Level Descriptors at Below Sufficient Level

Indicators	Unacceptable	Acceptable	Proficient
<u>Develops</u> learning objectives which are appropriate for the subject and grade level that are connected to the standards.	None or few of the learning objectives are appropriate to subject level <b>or</b> none or few of the learning objectives are connected to standards.	Most of the learning objectives are appropriate to subject level <b>and</b> most of the learning objectives are connected to standards.	All of the learning objectives are appropriate to subject level, are clearly connected to standards, and accommodate learning differences.
<u>Plans</u> appropriate and logically sequenced instructional strategies.	Planned instructional strategies are inappropriate or illogical.	Planned instructional strategies are appropriate and logical.	Planned instructional strategies are appropriate to subject and developmental levels and are clearly logical in sequence.

## Scoring Example 2:

# Proficiency Level Descriptors at Sufficient Level

Indicators	Unacceptable	Acceptable	Proficient
<p><b>1. <u>Develops</u> learning objectives which are appropriate for the subject and grade level, and are connected to the standards.</b> CAEP 1.1; InTASC 7</p>	<p>Objectives are inappropriate for the subject area/developmental level of learners <b>by being either too difficult or too easy for students</b>. Objectives are not stated in measurable terms, do not include criteria, and/or are not connected to the standards.</p>	<p>Objectives are appropriate for subject area/developmental level of learners, are connected to the standards, <b>and provide appropriate challenges for students (tasks are neither too easy nor too difficult)</b>. Objectives are measurable.</p>	<p>Objectives are appropriate for the subject area/developmental level of learners, are connected to the standards., provide appropriate challenges, <b>and incorporate multiple domains of learning</b>. Objectives are measurable and <b>each contains criteria for student mastery</b>.</p>
<p><b>2. <u>Plans</u> appropriate and logically sequenced instructional strategies.</b> CAEP 1.1; InTASC 7</p>	<p>Instructional strategies are incongruent with objectives. Majority of strategies are developmentally inappropriate. <b>Learning/practice tasks are arranged randomly in the lesson with steps between progressions either too large or too small to facilitate skill mastery.</b></p>	<p>Majority of instructional strategies are congruent with objectives, instructional strategies are developmentally appropriate, and lesson sequence is logical. <b>Learning/practice tasks are arranged in sequential and progressive steps to facilitate learning.</b></p>	<p>Instructional strategies are congruent with objectives. All strategies are developmentally appropriate and <b>address varied student needs</b>. <b>Learning/ practice tasks allow students to begin and end at different levels based on individual readiness and include progressive opportunities for students to extend tasks to increase the challenge.</b></p>



# Tips for Scoring and Proficiency Level Descriptors

Consider the following:

- Use an actual range of artifacts or an actual set of videos to define levels in performance terms.
- Determine what constitutes moving down a level or up a level.
- **Bold** critical aspects of performance.

Beware the following:

- Words ending in –LY (consistently, frequently, etc.)
- Wiggle words (sometimes, often, never, always)
- Frequency counts (fewer than 2, 3, 4 or more)
- Using the language of the standard at each performance level.
- Misalignment of an indicator or the “spirit” of an indicator with the scoring levels.

# Task Workshop

\*Worksheet #3: Scoring: Proficiency Level Descriptors (PLDs)

## **Task:**

- Using the assessment on Handout #1, circle language that is too subjective, may be hindering actionable feedback, is not performance-based, or is not a developmental sequence.
- Select one indicator's PLDs on Handout #1 in need of improvement and revise using the checklist on Worksheet #3.
- Time permitting, make a general list of the types of revisions needed for Handout #1.





# Next Steps and Questions

# ***Development and Analysis of EPP-Created Assessments***

## ***Worksheet and Handout Packet***

### **Table of Contents**

Handout #1: COEHS Impact on P-12 Student Learning Assessment Guidelines and Rubric	2
Worksheet #1: Administration and Purpose.....	5
Worksheet #2: Content of Assessment: Indicators.....	6
Worksheet #3: Scoring: Proficiency Level Descriptors (PLDs).....	7

*Based on the workshop offered by CAEP and GaPSC in Atlanta, GA on April 18, 2017*

Special Thanks to Margaret D. Crutchfield, CAEP Consultant; Deborah Eldridge, Consultant, LCV, Inc.; Brian Beitzel, SUNY Oneonta; Maureen Linton, Consultant, UCA Consulting, LLC; Nancy Mundshenck, Southern Illinois University, Carbondale; Diane Penland, Walden University; GaPSC



# Handout 1: COEHS Impact on P-12 Student Learning Assessment Guidelines and Rubric

## COEHS Impact on P-12 Student Learning Assessment Guidelines

The COEHS Impact on P-12 Student Learning Assessment is completed by all initial teacher candidates during clinical practice (student teaching or internship). This assessment will be completed as part of a sequential unit of instruction that may range from a few lessons in one week to several lessons over a course of more than one week. This project may be an added component of your edTPA learning segment or a completely different set of sequential lessons. This project should be part of the planned grade level curriculum in the P-12 classroom in which you are completing clinical practice. In other words, it should be part of what you would already be expected to teach. You will develop a pre- and post-assessment that will be administered to your P-12 students. Be sure to consult with your P-12 mentor teacher to determine which unit of study you should plan to teach and gain the approval of your mentor and university supervisor prior to beginning this project.

The purpose of this assessment is to collect and analyze assessment data on the P-12 students you are teaching during clinical practice to determine whether or not you are positively impacting their learning and use these data to make decisions about future instruction and remediation plans for improving the achievement of students.

The following components are required for this project:

1. **Unit Lesson Plans**: Lesson plans for each day of instruction in the unit. The lesson plans should follow the lesson plan format required by your program.
2. **Description of the Context for Learning**: Follow the guidelines provided for your program.
3. **Pre-assessment**: Conduct a pre-assessment to determine your students' knowledge of the central focus and/or learning objectives you are planning to teach in this unit.
  - a) Include a copy of the pre-assessment.
  - b) Include a display of data collected (e.g., tables, graphs, spreadsheets) from the pre-assessment. Follow the guidelines approved for your program.
  - c) Provide an analysis of what you learned from the pre-assessment data and how you used this information to modify instruction.
4. **Post-assessment**: Include a copy of the post-assessment instrument you administered at the end of the instructional unit if it is different from the pre-assessment.
5. **Data Analysis**: Analyze the data for the whole class AND sub-groups OR individual levels.
  - a. **Whole Class**:

Compile and display the data (e.g., tables, graphs, spreadsheets) for the entire class. Do not use actual names of the students. The display should show the comparison of the pre-and post-assessment results for each objective/outcome. Describe the data using simple descriptive techniques including range, mean, median, and mode.

**b. Subgroups or Individuals:**

Subgroups - Compile and display the data from at least one learning objective/outcome into groups for comparison based on at least one relevant factor (e.g., ethnicity, gender, language, race). Compare and contrast the subgroup data. Include specific examples for each subgroup.

**or**

Individuals - Select at least three students who represent different levels of performance. Analyze the assessment data you have gathered for each student and describe the specific performance of these individual students related to the learning objectives/outcomes.

6. **Reflection of Student Learning and Teaching**: Reflect on your performance as a teacher. Discuss your teaching in regard to the student learning results and professional development goals. Provide an overall analysis that describes what you gleaned from the pre-assessment, formative assessment, and post-assessment results. Describe what learning took place from the pre- to post-assessment. In what areas did students perform well? In what areas did the students perform less well? What do you think contributed to these differences? How did you use formative assessment to modify instruction? Based on this reflection of student learning and your teaching, what are your goals for future professional develop?



## COEHS Impact on P-12 Student Learning Assessment Rubric

Indicator	Level 4 Indicator Exceeded	Level 3 Indicator Met	Level 2 Indicator Partially Met	Level Indicator Not Met
Clarity and Accuracy of Evidence	Evidence is clearly organized, easy to understand, and contains no errors of representation. Data are comprehensive, including raw scores, mastery for each learning objective, summarized scores, and appropriate graphs.	Evidence is easy to understand and contains no errors of representation.	Evidence is understandable and contains few errors.	Evidence is not clear and accurate; it does not accurately reflect the data.
Alignment with Learning Objectives	Analysis is fully aligned with learning objectives and provides a comprehensive profile of student learning for the whole class and subgroups or individuals. Analysis shows clear evidence of understanding which students learned mastered each objective.	Analysis is fully aligned with learning objectives and provides a comprehensive profile of student learning for the whole class and subgroups or individuals.	Analysis of student learning is partially aligned with learning objectives and/or fails to provide a comprehensive profile of student learning relative to the learning objectives/outcomes for the whole class and subgroups or individuals.	Analysis of student learning is not aligned with learning objectives.
Interpretation of Data	Interpretation of data is meaningful, and appropriate conclusions are drawn from the data. Analysis is thorough and supported by appropriate tables and graphs.	Interpretation of data is meaningful, and appropriate conclusions are drawn from the data.	Interpretation of data is technically accurate, but conclusions are missing or not fully supported by data.	Interpretation of data is inaccurate, and conclusions are missing or unsupported by data.
Evidence of Impact on Student Learning	Analysis of student learning includes evidence of the impact on student learning in terms of numbers of students who achieved and made progress toward each learning objective.	Analysis of student learning includes evidence of the impact on student learning in terms of numbers of students who achieved and made progress toward the objectives.	Analysis of student learning includes incomplete evidence of the impact on student learning in terms of numbers of students who achieved and made progress toward the objectives.	Analysis of student learning fails to include evidence of impact on student learning in terms of numbers of students who achieved and made progress toward objective.
Reflection	Reflection is relevant and well developed. It relates directly to the student learning assessment data and specific goals for future professional development.	Reflection is relevant and includes some reference to student learning and specific goals for future professional development.	Reflection is ancillary and minimally references student learning and specific goals for future professional development.	Reflection fails to relate directly to the student learning assessment data. and/or Specific goals for future professional development were not appropriate.

# Worksheet 1: Administration and Purpose

## Task:

1. Draft a context statement using the assessment on Handout #1.

## In your statement, be sure to answer the following questions:

1. What are the point(s) in the program when the assessment is administered?
2. How often is it administered?
3. Who administers it at each different point?
4. What is the purpose of this assessment (e.g., what aspects of candidate performance are you measuring?)
5. How is this assessment used to make decisions about candidates' progress through the program?
6. What score/rating, etc. must a candidate reach in order to pass this assessment?
7. What happens if a candidate does not meet this required criterion?
8. Is there a discussion of the alignment of this assessment and indicators to standards? How was this alignment done?



## Worksheet 2: Content of Assessment: Indicators

**Tasks:**

1. Select at least two indicators from the assessment on Handout #1 that do not currently meet the sufficient Level.
2. Rewrite the indicators following the guidelines provided on the worksheet.
3. Time permitting, evaluate the other indicators and make a list of needed revisions.

**For Worksheet 2, answer the following questions for each indicator:**

Question:	Indicator 1	Indicator 2
Is the indicator tagged to CAEP, InTASC, national/ professional and state standards (as appropriate)?		
Is the standard(s) cited appropriate for the indicator?		
Are the indicators written at an appropriate level of difficulty or level of effort as described in the standards?		
Are the indicators unambiguous? Do they clearly describe the candidate proficiencies to be evaluated?		
If the standard that is cited addresses higher level functioning, does the indicator require higher levels of intellectual behavior?		
What percentage of indicators in the assessment address consequential attributes of candidate proficiency?		

## Worksheet 3: Scoring: Proficiency Level Descriptors (PLDs)

**Tasks:**

1. Using the assessment on Handout #1, circle language that is too subjective, may be hindering actionable feedback, is not performance-based, or is not a developmental sequence.
2. Select one indicator's PLDs in need of improvement and revise using the checklist on the worksheet.
3. Time permitting, make a general list of the types of revisions needed for your assessment.

**For Worksheet 3, complete the following checklist for each PLD:**

**Indicator Number:** \_\_\_\_\_

CAEP Sufficient Level	Currently in Rubric (Y/N):			Revision(s) Needed
	PLD 1	PLD 2	PLD 3	
Basis for judging candidate performance is well defined.				
Each PLD is qualitatively defined.				
PLDs represent a developmental sequence from level to level.				
Feedback provided to candidates is actionable.				
PLDs are defined in actionable, performance-based, or observable behavior terms.				
<i>Are wobble words used? (never, sometimes, often, consistently, always)</i>				
<i>Are -LY words used (consistently, frequently, clearly, accurately)?</i>				
<i>Are frequency counts used? (1 or 2, 3 or more)</i>				
<i>Is the language of the standard used repeatedly?</i>				