Watermark Assessment
Request Form

Protocol: The Council for the Accreditation of Educator Preparation (CAEP) *Evaluation Framework for Education Preparation Provider (EPP) Created Assessments* serves as the foundation for this request form and protocol. Please complete this form for EACH assessment (including subject and pedagogical content tests, observations, projects, assignments, and rubrics) to be submitted in Watermark. This would be applicable to a New assessment or Revised assessment. The New or Revised assessment should be submitted at least 30 days prior to the Fall semester in which it is to be implemented. (unless the course is only offered once a year in Spring or Summer which will need to be requested at least one cycle/semester before implementation).

Prior to completing the Watermark Assessment Request form, please ensure the cover page of your rubric includes the following:
 1) Course Title, 2) Course Number, 3) Standards, 4) Purpose, 5) Administration, and 6) Success Indicator.

Additionally, on the footer of your assessment/rubric, list the semester and year of the revision and addition. Please make sure to include the directions/instructions of the assessment with the rubric or grading instrument along with this completed Watermark Rubric Form.

The faculty member who is requesting the New or Revised assessment in Watermark may reach out to their Departmental Contact (see below) or the Assistant Dean, Dr. Mitzy Johnson for assistance in completion of this form. There is also an approval required from the Program Coordinators from all campuses prior to the submission and supporting documentation necessitating the addition or revision of the assessment.

Watermark Contacts are as follows:

* Curriculum, Instruction, & Special Education-Ms. Cindy Stevenson
* Music & Physical Education-Ms. Jenny Hartness
* Undergraduate Teaching Internship-Ms. Jennifer Neilson
* Alternate Route programs-Stephanie Etheridge & Elizabeth Palmer
* Educational Leadership-Dr. Katie Oswalt

* All Counseling, Educational Psychology, and Foundations-Ms. Amy Stockton

TIPS for Writing Appropriate Scoring Level in your Rubric

Consider the following:

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

BEWARE of 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**

Example of Rubric Cover Page

Title of the Assessment: TEACHER INTERN ASSESSMENT INSTRUMENT (TIAI)

Semester Added or Revised: Fall 2018

Course Title:

Course Number:

Standards: (List the accreditation name/standard date and actual ones being tagged on the rubric/assessment)

Purpose: To provide a comprehensive assessment (both formative and summative) of the teaching practice of teacher candidates.

Administration: This instrument is administered by classroom mentor teachers and university supervisors, formative and summative, during each field experience placement in _____ (list the course title and number).

Success Indicator: Items rated at the "Meets Standard" level represent successful teaching practice by the teacher candidate. Anything below "Meets Standard" can be seen as an area in need of improvement.

List/Tag the Standards under each Indicator)

Indicators	Unacceptable (0)	Needs Improvement (1)	Meets Standard (2)	Exceeds Standard (3)
1. Develops measurable and observable grade and subject level objectives that are aligned with appropriate state curricula frameworks. **(CAEP 1.2, InTASC 7, TGR 1)**	Objectives are not measurable, observable, or aligned with appropriate state curricula frameworks.	Objectives are aligned with appropriate state curricula frameworks, but they are not measurable or observable.	Objectives are measurable, observable, and aligned with appropriate state curricula frameworks.	**In addition to meets standard**, objectives are stated at different instructional levels based on individual needs of students (DOK Levels and/or Bloom's Taxonomy).

Please complete the fields within the Watermark Request Form below:

(Revised Spring 2023)

Submission Date:	
Department:	
Program and degree level:	
Concentration Area:	
Course Title and Number:	
Please check if a New Assessment OR Replace Existing Assessment	☐ **New Assessment** ☐ **Replace Existing Assessment** **If replacing existing assessment, please list the name of the old assessment below:**
Title of the New or Revised Assessment	
Type of Assessment of the New or Revised Assessment	☐ **Rubric** ☐ **Test** ☐ **Project** ☐ **Assignment** ☐ **Observation** ☐ **Other**_____ **(please explain)**
Evaluation Method, please check the appropriate method:	☐ **Rubric (refer to Section III: Scoring)** ☐ **Meets Requirement/Does Not Meet Requirement** (This is not used often.)
Faculty member completing this form:	
Please list the names of faculty members involved in the development of the assessment or revision:	
Nature of Change and what precipitated the Change:	

Revised Spring 2023)

Provide supporting documentation (Please include a copy of the syllabus and attach directions of the assignment as well as evidence of faculty collaboration such as attach minutes of program faculty meeting, advisory board minutes, or email, etc.):	
Program Coordinator(s) Approval for all Campuses along with Date of Approval	
Watermark Department Contact Acknowledgement/Review of the request as well as Assistant Dean Review (Including Date of Review and notes). For your reference, the Watermark Department Contacts are as follows: • Curriculum, Instruction, & Special Education-Ms. Cindy Stevenson • Music & Physical Education-Ms. Jenny Hartness • Undergraduate Teaching Internship-Ms. Jennifer Neilson • Alternate Route programs-Stephanie Etheridge & Elizabeth Palmer • Educational Leadership-Dr. Katie Oswalt • All Counseling, Educational Psychology, and Foundations-Ms. Amy Stockton	
Assistant Dean Signature/Date of Review:	
Fall Semester Requested for Change (unless the course is only offered once a year in Spring or Summer which will need to be requested at least one cycle before implementation):	

Revised Spring 2023)

I. ADMINISTRATION AND PURPOSE

1. Explain the purpose of this assessment (e.g. what aspects of the College of Education (COE) students' performance area are you measuring) and how is this assessment used to make decisions about the COE students' progress through the program?

2. Is the overall purpose and administration listed on the assessment and directions?　Yes ☐　　No ☐

3. At what point or points in the program of study is this assessment administered? (e.g. first year, last year, entry course, exit course, etc.)?

4. Are instructions provided to COE students to complete the assessment informative and unambiguous?
Yes ☐ No ☐

5. Is the judgment (criterion for success, or what is "good enough") made explicit for COE students and identified as the Success Indicator on the assessment and directions?
Yes ☐ No ☐

6. **Are the evaluation categories/indicators on the assessment tagged with CAEP, InTASC, national/professional (SPA) and state standards? (Note: Be careful not to double dip standards to one criterion (e.g. InTASC 3 &4)**

Yes ☐ No ☐

List the standards and version (year) tagged below: (Ex. CAEP, CACREP, NCTM, NCTE, etc.) For CACREP, if Key Performance Indicator (KPI) is tagged, please ensure to note it with the appropriate abbreviation and number: CORE KPI _, CMH KPI _, Rehab KPI _, SC KPI _, and Doctoral KPI _

```



```

(If NO is selected for any of the above responses for Section I, please revise the assessment to meet this requirement before it is submitted.)

II. CONTENT OF ASSESSMENT

(The word "Indicator" (aka Criterion) is used as a generic term for the assessment items. For content tests, the term refers to a question. For projects or assignments, it refers to a prompt or task that the COE student is to perform. For an observation, an indicator might be a category of performance to observe or a specific aspect of COE student's performance that a reviewer would record.

1. **The indicators on the assessment explicitly identified aspects of the CAEP, InTASC, national/professional and state standards?** Yes ☐ No ☐

2. **The indicators are congruent with the complexity, cognitive demands and skill requirements described in the standards (e.g. create, evaluate, analyze, & apply). For example, when a standard requires the COE students to "demonstrate" problem solving, then the indicator is specific to the COE students' application of knowledge to solve problems.** Yes ☐ No ☐

3. **The level of COE student's effort required, or the difficulty or degree of challenge of the assessment is consistent with the standards and is reasonable for the COE students who are making appropriate progress toward being ready to teach or take on the other professional responsibilities.** Yes ☐ No ☐

4. **For reliability, most indicators (at least those comprising 80% of the total score) require observers to judge consequential attributes of COE student proficiency.** Yes ☐ No ☐

(If NO is selected for any of the above responses in Section II, please revise the assessment to meet this requirement before it is submitted.)

III. SCORING for RUBRIC

1. **How many levels are represented on the scale of the assessment, not including "no data" or "unobserved" category?**

   (Please note: It is highly encouraged that "no data" or "unobserved categories" not be used and there must be at least three well-developed levels. It is recommended specifically 4 levels)

   (Check the number below):

   ☐   1 (If 1 is selected, please revised the assessment to meet this requirement)

   ☐   2 (If 2 is selected, please revised the assessment to meet this requirement)

   ☐   3

   ☐   4

   ☐   5

   ☐   Other: [   ]

2. **Is each of the performance level descriptors qualitatively defined by specific criteria aligned with indicators of the assessment?** Yes ☐ No ☐
   List the performance level descriptors below that correlates with the scale

   [   ]

3. **The performance levels represent developmental sequence from level to level (to provide raters with explicit guidelines for evaluating COE student performance and for providing COE students' with explicit feedback on their performance).**
Yes ☐ No ☐

4. **Performance level attributes are defined in actionable, performance-based, or observable terms. (e.g. Bloom's, etc.)**
Yes ☐ No ☐

5. **The COE student's final score on the assessment is clearly explained and is fair and reasonable and the same across all campuses.** Yes ☐ No ☐

   List the total points of the assessment below:

   List the total points that are listed on the syllabus for this assessment below:

6. **Feedback provided to COE students is actionable for it is directly related to the preparation of the program and can be used for program improvement as well as feedback to the COE student.** Yes ☐ No ☐

(If NO is selected for any of the above responses in Section III, please revise the assessment to meet this requirement before it is submitted.)

Content Validity and Reliability

If the assessment is not Proprietary (Statewide Common Assessment, Specialized Professional Association Designed Rubric, etc.) Content Validity (CV) and Reliability of the assessment instrument must be established. Refer to Guidelines for establishing Content Validity and Reliability in this document.
Please mark one below.

Statewide Assessment ☐ SPA Designed Assessment ☐ Program Designed Assessment (EPP) ☐

NOTES

For Program Designed Assessment (EPP), after using the Guidelines for establishing Content Validity and Reliability, please complete the information below:

-CVR for this assessment:
-Names of Panelists were:
-Calculation of the Percentage of Inter-rater agreement:
-List the Date and names of scorers of when the inter-rater agreement occurred:

Guidelines for Establishing Content Validity

The *CAEP Evaluation Framework for Education Preparation Provider (EPP) Created Assessments* (Section 5) requires EPPs to demonstrate the validity of all EPP created assessment instruments (with the exception of surveys). At minimum, EPPs are required to report how they have established **content validity**. This refers to the appropriateness of the content of an assessment; does it measure/assess what COE students must demonstrate. At the heart of content validity requires judgement of experts. Please note that aligning an assessment instrument with professional standards does not establish content validity!

Lawshe's (1975) quantitative approach to content validation has been cited by CAEP as the useful method for EPPs to follow. The steps below summarize Lawshe's process.

Protocol:

Step 1:   Assemble a pool of potential items to be included on the instrument
Note: If this instrument is to be deployed in a clinical setting (e.g. public schools) be aware that you should require personnel from that context to be involved in the construction of the assessment instruments.

Step 2:   Subject the items to expert review
a. Present each individual item to each member of the review panel
(at least 3 to 5 should be on the panel)

b. Along with the item, include a forced-choice response asking if the skill/knowledge measured by the item is:

- Essential
- Useful but not essential
- Not necessary
... to the target evaluation category/criterion (e.g student intern performance, etc.)

Step 3:   Follow the guidance below to determine which items should be included on the instrument.

Content Validity Ratio (CVR) Calculation:

In order for an item to be included, Lawshe (1975) recommended that at least half of the expert panel agree that the item represents a characteristic that is essential to the criterion task/behavior. Lawshe introduced a metric called Content Validity Ration (CVR):

$$CVR = \frac{ne - N/2}{N/2}$$

Where ne = number of panelists who rate the item as Essential
N = total number of panelists
CVR for this assessment is _____
Panelists were: _____

Guidelines for Establishing Reliability

(Revised Spring 2023)

The *CAEP Evaluation Framework for Education Preparation Provider (EPP) Created Assessments* (CEF, Section 4) requires EPPs to demonstrate reliability of data resulting from all EPP created assessment instruments (with the exception of surveys). At minimum, EPPs are required to report the **percentage of inter-rater agreement.** As with validity, the obligation of EPPs pertaining to reliability is to employ "accepted research standards,"-in other words, approaches that would survive the scrutiny of the peer-review process for a manuscript submitted to a journal.

There are two important components to meeting CAEP's expectation for reliability:
1) Training raters to score consistently
2) Verifying the consistency of raters' scores

Scorer Training and Calibration

CEF 4.b. requires EPPs to report how they have trained scorers so that everyone who uses the instrument scores consistently. The following steps may be helpful in guiding this process.

Step 1: Assemble a pool of samples to be evaluated using the instrument. These could be documents (e.g. student-authored papers), videos, (student interns), etc.

Step 2: For each sample, arrive at a consensus score for each indicator on the instrument.

Step 3: Host scorer training sessions. Walk new scorers through a few samples (from Step 1) and explain why the consensus scores (from Step 2) are the correct ones. Then, use some of the other samples (from Step 1) for scorers to independently practice; then compare their scores to the consensus scores (from Step 2) for those samples.

Step 4: Repeat Step 3 as needed, immediately and over time, until satisfactory consistency (see the second row in the table below; the consensus scores for the sample being evaluated will be one "scorer") is achieved for both new and experienced scorers.

Checking on Inter-Rater Agreement and Reliability

CEF 4.b requires EPPs to report the results of their checks on inter-rater agreement and reliability. There are several approaches that can be used to document scorer consistency.

Number of Scorers Tested	Potential Approaches
One	Because CAEP requires scorer *agreement*, at least one additional scorer will need to be trained (see above guidelines for scorer training).
Two	*Option 1:* Calculate the percentage of inter-rater agreement (i.e. the proportion of indicators scored identically); also consider including (as a comparison) the percentage of agreement within one point. *Option 2:* Calculate the reliability coefficient (i.e. the correlation of scores from the two scorers).
More than two	*Option 1:* Calculate an agreement statistic (e.g. kappa) *Option 2:* Calculate a reliability coefficient (e.g. ICC, Krippendorff)

Note: Data from one sample WILL NOT be sufficient to satisfactorily demonstrate consistency among scorers.

List the Calculation of the Percentage of Inter-rater agreement: _____ and date and names of scorers of when then inter-rater agreement occurred:_____

References

Ayre, C. & Scally, A.J (2014). Critical values of Lawshe's content validity ratio:

(Revised Spring 2023)

Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development,* 47, 79-86.

Lawshe, C.H. (1975). A quantitative approach to content validity. *Personnel Psychology,* 28, 563- 575.

Eldridge, D & Crutchfield, M. (Spring 2017). *Assessment Acclamation Improvement Together CAEP Pre-conference Workshop*

Rivadeneyra, T. (Fall 2017) *Establishing Validity and Reliability for EPP-Created Assessments CAEP Post-conference Workshop*

Acknowledgement: Special thanks to University of North Carolina-Charlotte's willingness to share their *Initial Rubric Review Form*

(Revised Spring 2023)