

Resources for Faculty

Revising SLOs and Rubrics to Assess SLOs

1. The College of Education Office of Educational Assessment – we are happy to serve as a resource to faculty as we collaboratively complete this work. Please let Laura or Ashley know if we can help, particularly when considering rubric revisions. Laura is planning to attend as many of the program/dept meetings as possible and Ashley can answer many questions about what is possible in Taskstream (rubrics vs. forms, etc.). We are also available for individual meetings.

Dr. Laura Hart: lchart1@uncc.edu

Ashley Flatley: aflatley@uncc.edu

Office phone for both: 704-687-8163

2. Included documents: Several documents are included as appendices. These documents are intended to support and inform faculty decision-making through this process.
 - a. Appendix A – CAEP Standards for Initial and Advanced Licensure Programs
 - b. Appendix B – Guidelines from CAEP: Developing Assessment Instruments
 - c. Appendix C – CAEP draft: Rubrics for Evaluation of EPP Instruments Used as Accreditation Evidence

Standards for Initial and Advanced Programs

as approved by the CAEP Board of Directors

June 5, 2014

Definitions:

Initial Programs. Programs at baccalaureate level (undergraduate) or graduate certificate level for candidates who are seeking their initial, first teaching license. Phase I of the Graduate Certificate program leads to initial teacher licensure. Phase I candidates are considered “initial” candidates, even though they are enrolled in graduate-level coursework.

Phase II of graduate certificate programs may lead to an MAT degree, but this is optional. Phase II candidates who opt to continue in the MAT program are considered “advanced” candidates, as completion of the MAT does lead to advanced teacher licensure.

Advanced Programs. Programs at post-baccalaureate levels for (1) the continuing education of teachers who have previously completed initial preparation or (2) the preparation of other school professionals. Advanced programs commonly award graduate credit and include master’s, specialist, and doctoral degree programs as well as non-degree licensure programs offered at the post-baccalaureate level. Examples of these programs include:

- Teachers who are preparing for a second license at the graduate level in a field different from the field in which they had their first license;
- Programs for teachers who are seeking a master’s degree in the field in which they teach;
- Programs not tied to licensure, such as programs in curriculum and instruction; and
- Programs for other school professionals, such as school counselors, school psychologists, educational administrators, and reading specialists.

Note: In Standard 1, the subjects of components are “candidates.” The specific knowledge and skills described will develop over the course of the preparation program and may be assessed at any point, some near admission, others at key transitions such as entry to clinical experiences and still others near candidate exit as preparation is completed.

Completer refers to any candidate who is exiting a preparation program by successfully satisfying the requirements of the Educator Preparation Provider (EPP).

Standard 1

The provider ensures that candidates develop a deep understanding of the critical concepts and principles of their discipline and, by completion, are able to use discipline-specific practices flexibly to advance the learning of all students toward attainment of college- and career-readiness standards.

Initial Components**Advanced Program Components****Candidate Knowledge, Skills, and Professional Dispositions**

1.1 Candidates demonstrate an understanding of the 10 InTASC standards at the appropriate progression level(s)^[i] in the following categories: the learner and learning; content; instructional practice; and professional responsibility.

Candidate Knowledge, Skills, and Professional Dispositions

1.1 Advanced program candidates demonstrate an understanding and are able to apply knowledge and skills specific to their discipline.

Provider Responsibilities

1.2 Providers ensure that completers use research and evidence to develop an understanding of the teaching profession and use both to measure their P-12 students' progress and their own professional practice.

Provider Responsibilities

1.2 Providers ensure that advanced program completers use research and evidence to develop school environments that support and assess P-12 students' learning and their own professional practice specific to their discipline.

1.3 Providers ensure that completers apply content and pedagogical knowledge as reflected in outcome assessments in response to standards of Specialized Professional Associations (SPA), the National Board for Professional Teaching Standards (NBPTS), states, or other accrediting bodies (e.g., National Association of Schools of Music – NASM).

1.3 Providers ensure that advanced program completers apply content and discipline-specific knowledge as reflected in state and/or national discipline-specific standards where they exist including Specialized Professional Associations (SPAs) and other accrediting bodies (e.g., Council for Accreditation of Counseling and Related Educational Programs - CACREP).

1.4 Providers ensure that completers demonstrate skills and commitment that afford all P-12 students access to rigorous college- and career-ready standards (e.g., Next Generation Science Standards, National Career Readiness Certificate, Common Core State Standards).

1.4 Providers ensure that advanced program completers demonstrate skills and commitment to creating supportive environments that afford all P-12 students access to rigorous college- and career-ready standards (e.g., Next Generation Science Standards, National Career Readiness Certificate, Common Core State Standards).

1.5 Providers ensure that completers model and apply technology standards as they design, implement and assess learning experiences to engage students and improve learning; and enrich professional practice.

1.5 Providers ensure that advanced program completers model and apply technology standards as they design, implement and assess learning experiences/environments to engage students and improve learning; and enrich professional practice.

Standard 2

The provider ensures that effective partnerships and high-quality clinical practice are central to preparation so that candidates develop the knowledge, skills, and professional dispositions necessary to demonstrate positive impact on all P-12 students' learning and development.

Initial Components**Partnerships for Clinical Preparation**

2.1 Partners co-construct mutually beneficial P-12 school and community arrangements, including technology-based collaborations, for clinical preparation and share responsibility for continuous improvement of candidate preparation. Partnerships for clinical preparation can follow a range of forms, participants, and functions. They establish mutually agreeable expectations for candidate entry, preparation, and exit; ensure that theory and practice are linked; maintain coherence across clinical and academic components of preparation; and share accountability for candidate outcomes.

Clinical Educators

2.2 Partners co-select, prepare, evaluate, support, and retain high-quality clinical educators, both provider- and school-based, who demonstrate a positive impact on candidates' development and P-12 student learning and development. In collaboration with their partners, providers use multiple indicators and appropriate technology-based applications to establish, maintain, and refine criteria for selection, professional development, performance evaluation, continuous improvement, and retention of clinical educators in all clinical placement settings.

Advanced Program Components**Partnerships for Clinical Preparation**

2.1 Partners co-construct mutually beneficial P-12 school and community arrangements, including technology-based collaborations, for clinical preparation and share responsibility for continuous improvement of advanced program candidate preparation. Partnerships for clinical preparation can follow a range of forms, participants, and functions. They establish mutually agreeable expectations for advanced program candidate entry, preparation, and exit; ensure that theory and practice are linked; maintain coherence across clinical and academic components of preparation; and share accountability for advanced program candidate outcomes.

Clinical Educators

2.2 Partners co-select, prepare, evaluate, support, and retain high-quality clinical educators, both provider- and school-based, who demonstrate a positive impact on advanced program candidates' development and P-12 student learning and development. In collaboration with their partners, providers use multiple indicators and appropriate technology-based applications to establish, maintain, and refine criteria for selection, professional development, performance evaluation, continuous improvement, and retention of clinical educators in all clinical placement settings.

Clinical Experiences

2.3 The provider works with partners to design clinical experiences of sufficient depth, breadth, diversity, coherence, and duration to ensure that candidates demonstrate their developing effectiveness and positive impact on all students' learning and development. Clinical experiences, including technology-enhanced learning opportunities, are structured to have multiple performance-based assessments at key points within the program to demonstrate candidates' development of the knowledge, skills, and professional dispositions, as delineated in Standard 1, that are associated with a positive impact on the learning and development of all P-12 students.

Clinical Experiences

2.3 The provider works with partners to design clinical experiences of sufficient depth, breadth, diversity, coherence, and duration to ensure that advanced program candidates demonstrate their developing effectiveness in creating environments that support all students' learning and development. Clinical experiences, including technology-enhanced learning opportunities, are structured to have multiple performance-based assessments at key points within the program to demonstrate advanced program candidates' development of the knowledge, skills, and professional dispositions, as delineated in Standard 1, that are associated with creating a supportive school environment that results in a positive impact on the learning and development of all P-12 students.

CAEP Standards with Components	CAEP Standards for Advanced Programs with Components
<p>Standard 3 <i>The provider demonstrates that the quality of candidates is a continuing and purposeful part of its responsibility from recruitment, at admission, through the progression of courses and clinical experiences, and to decisions that completers are prepared to teach effectively and are recommended for certification. The provider demonstrates that development of candidate quality is the goal of educator preparation in all phases of the program. This process is ultimately determined by a program’s meeting of Standard 4.</i></p>	<p>Standard 3* <i>The provider demonstrates that the quality of advanced program candidates is a continuing and purposeful part of its responsibility from recruitment, at admission, through the progression of courses and clinical experiences, and to decisions that advanced program completers are prepared to perform effectively and are recommended for certification where applicable. The provider demonstrates that development of candidate quality is the goal of educator preparation in all phases of the program. This process is ultimately determined by a program’s meeting of Standard 4.</i></p> <ul style="list-style-type: none"> • Change of wording of the standard is noted in red.
<p>Plan for Recruitment of Diverse Candidates who Meet Employment Needs 3.1 The provider presents plans and goals to recruit and support completion of high-quality candidates from a broad range of backgrounds and diverse populations to accomplish their mission. The admitted pool of candidates reflects the diversity of America’s P-12 students. The provider demonstrates efforts to know and address community, state, national, regional, or local needs for hard-to-staff schools and shortage fields, currently, STEM, English-language learning, and students with disabilities.</p>	<p>Plan for Recruitment of Diverse Candidates who Meet Employment Needs 3.1 The provider presents plans and goals to recruit and support completion of high-quality advanced program candidates from a broad range of backgrounds and diverse populations to accomplish their mission. The admitted pool of candidates reflects the diversity of America’s P-12 students. The provider demonstrates efforts to know and address community, state, national, regional, or local needs for school and district staff prepared in advanced fields.</p>
<p>Admission Standards Indicate That Candidates Have High Academic Achievement and Ability 3.2 The provider sets admissions requirements, including CAEP minimum criteria or the state’s minimum criteria, whichever are higher, and gathers data to monitor applicants and the selected pool of candidates. The provider ensures that the average grade point average of its accepted cohort of candidates meets or exceeds the</p>	<p>Admission Standards Indicate That Candidates Have High Academic Achievement and Ability 3.2 The provider sets admissions requirements, including CAEP minimum criteria, the state’s minimum criteria, or graduate school minimum criteria, whichever is highest, and gathers data to monitor applicants and the selected pool of candidates. The provider ensures that the average grade point average of its accepted cohort of candidates [meets or exceeds the CAEP minimum of 3.0, or the group average performance on</p>

<p>CAEP minimum of 3.0, and the group average performance on nationally normed ability/achievement assessments such as ACT, SAT, or GRE:</p> <ul style="list-style-type: none"> ▪ is in the top 50 percent from 2016-2017; ▪ is in the top 40 percent of the distribution from 2018-2019; and ▪ is in the top 33 percent of the distribution by 2020.[i] <p>If any state can meet the CAEP standards, as specified above, by demonstrating a correspondence in scores between the state-normed assessments and nationally normed ability/achievement assessments, then educator preparation providers from that state will be able to utilize their state assessments until 2020. CAEP will work with states through this transition.</p> <p>Over time, a program may develop a reliable, valid model that uses admissions criteria other than those stated in this standard. In this case, the admitted cohort group mean on these criteria must meet or exceed the standard that has been shown to positively correlate with measures of P-12 student learning and development.</p> <p>The provider demonstrates that the standard for high academic achievement and ability is met through multiple evaluations and sources of evidence. The provider reports the mean and standard deviation for the group.</p>	<p>a nationally normed ability/achievement assessment such as GRE, MAT, or other required graduate level assessment is in the top 50%.]</p> <p>Over time, a program may develop a reliable, valid model that uses admissions criteria other than those stated in this standard. In this case, the admitted cohort group mean on these criteria must meet or exceed the standard that has been shown to positively correlate with measures of supportive environments that positively impact on all P-12 student learning and development.</p> <p>The provider demonstrates that the standard for high academic achievement and ability is met through multiple evaluations and sources of evidence. The provider reports the mean and standard deviation for the group.</p>
<p>Additional Selectivity Factors</p> <p>3.3 Educator preparation providers establish and monitor attributes and dispositions beyond academic ability that candidates must demonstrate at admissions and during the program. The provider selects criteria, describes the measures used and evidence of the reliability and validity of those measures, and reports data that show</p>	<p>Additional Selectivity Factors</p> <p>3.3 Educator preparation providers establish and monitor attributes and dispositions beyond academic ability that advanced program candidates must demonstrate at admissions and during the program. The provider selects criteria, describes the measures used and evidence of the reliability and validity of those measures, and reports data that show how</p>

<p>how the academic and non-academic factors predict candidate performance in the program and effective teaching.</p>	<p>the academic and non-academic factors predict advanced program candidate performance in the program and in service.</p>
<p>Selectivity During Preparation 3.4 The provider creates criteria for program progression and monitors candidates' advancement from admissions through completion. All candidates demonstrate the ability to teach to college- and career-ready standards. Providers present multiple forms of evidence to indicate candidates' developing content knowledge, pedagogical content knowledge, pedagogical skills, and the integration of technology in all of these domains.</p>	<p>Selectivity During Preparation 3.4 The provider creates criteria for program progression and monitors candidates' advancement from admissions through completion. All advanced program candidates demonstrate the ability to create and maintain supportive environments for teaching college- and career-ready standards. Providers present multiple forms of evidence to indicate advanced program candidates' application of content knowledge and research, data-driven decision making, and the integration of technology in all of these domains.</p>
<p>Selection At Completion 3.5 Before the provider recommends any completing candidate for licensure or certification, it documents that the candidate has reached a high standard for content knowledge in the fields where certification is sought and can teach effectively with positive impacts on P-12 student learning and development.</p>	<p>Selection At Completion 3.5 Before the provider recommends any advanced program candidate for program completion, it documents that the advanced program candidate has reached a high standard for content knowledge; data- and research-driven decision making; and integration of technology in the discipline; and demonstrates the ability to create, maintain, and enhance supportive environments for effective P-12 learning.</p>
<p>3.6 Before the provider recommends any completing candidate for licensure or certification, it documents that the candidate understands the expectations of the profession, including codes of ethics, professional standards of practice, and relevant laws and policies. CAEP monitors the development of measures that assess candidates' success and revises standards in light of new results.</p>	<p>3.6 Before the provider recommends any advanced program candidate for program completion, it documents that the advanced program candidate understands the expectations of the profession, including codes of ethics, professional standards of practice, and relevant laws and policies. CAEP monitors the development of measures that assess candidates' success and revises standards in light of new results.</p>

*This document is intended to adapt the 3.2 admissions criteria for initial preparation to graduate level advanced preparation programs. The first sentence would parallel the provisions for initial preparation relevant to EPP admission requirements, a CAEP minimum, normed ability/achievement assessments, and monitoring the results for the admitted candidates. The bracketed phrase would maintain the 3.0 GPA. In this case there appears to be no nationally representative data, but available statistics suggest that 3.0 is in range of current GPA for college BA level work. The "top half" would be set as a criterion for cohort performance on a normed test of ability/achievement. However, there would be no phase-in period to a higher criterion (moving up to the top 40% and then the top 33%) as there is for initial preparation. Instead, CAEP could evaluate how the 50% level works in actual practice. Currently, for GRE verbal, the "top half" of all test takers who indicate their intended field of graduate study demonstrate similar performance for education as for engineering, physical sciences, life sciences and business. The normed test and GPA requirements would be alternatives (rather than additive) because current admissions criteria vary across institutions and individual graduate programs.

Standard 4

The provider demonstrates the impact of its completers on P-12 student learning and development, classroom instruction, and schools, and the satisfaction of its completers with the relevance and effectiveness of their preparation.

Impact on P-12 Student Learning and Development

4.1 The provider documents, using multiple measures that program completers contribute to an expected level of student-learning growth. Multiple measures shall include all available growth measures (including value-added measures, student-growth percentiles, and student learning and development objectives) required by the state for its teachers and available to educator preparation providers, other state-supported P-12 impact measures, and any other measures employed by the provider.

Impact on P-12 Student Learning and Development

4.1 The provider documents, using multiple measures, that advanced program completers create a supportive learning environment that contributes to an expected level of P-12 student-learning growth. Multiple direct and indirect measures shall include all available growth measures appropriate to the discipline, required by the state and available to educator preparation providers, other state-supported P-12 impact measures where applicable, and other measures employed by the provider.

Indicators of Teaching Effectiveness

4.2 The provider demonstrates, through structured and validated observation instruments and student surveys, that completers effectively apply the professional knowledge, skills, and dispositions that the preparation experiences were designed to achieve.

Indicators of Effectiveness

4.2 The provider demonstrates, through multiple measures, that advanced program completers effectively apply the professional knowledge, skills, and dispositions that the preparation experiences were designed to achieve.

Satisfaction of Employers

4.3. The provider demonstrates, using measures that result in valid and reliable data and including employment milestones such as promotion and retention, that employers are satisfied with the completers' preparation for their assigned responsibilities in working with P-12 students.

Satisfaction of Employers

4.3. The provider demonstrates, using measures that result in valid and reliable data and including employment milestones such as promotion and retention, that employers are satisfied with the advanced program completers' preparation for their assigned responsibilities in working with P-12 students.

Satisfaction of Completers

4.4 The provider demonstrates, using measures that result in valid and reliable data, that program completers perceive their preparation as relevant to the responsibilities they confront on the job, and that the preparation was effective.

Satisfaction of Completers

4.4 The provider demonstrates, using measures that result in valid and reliable data, that advanced program completers perceive their preparation as relevant to the responsibilities they confront on the job, and that the preparation was effective.

Standard 5

The provider maintains a quality assurance system comprised of valid data from multiple measures, including evidence of candidates' and completers' positive impact on P-12 student learning and development. The provider supports continuous improvement that is sustained and evidence-based, and that evaluates the effectiveness of its completers. The provider uses the results of inquiry and data collection to establish priorities, enhance program elements and capacity, and test innovations to improve completers' impact on P-12 student learning and development.

Quality and Strategic Evaluation

5.1 The provider's quality assurance system is comprised of multiple measures that can monitor candidate progress, completer achievements, and provider operational effectiveness. Evidence demonstrates that the provider satisfies all CAEP standards.

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.

Continuous Improvement

5.3 The provider regularly and systematically assesses performance against its goals and relevant standards, tracks results over time, tests innovations and the effects of selection criteria on subsequent progress and completion, and uses results to improve program elements and processes.

5.4 Measures of completer impact, including available outcome data on P-12 student growth, are summarized, externally benchmarked, analyzed, shared widely, and acted upon in decision-making related to programs, resource allocation, and future direction.

5.5 The provider assures that appropriate stakeholders, including alumni, employers, practitioners, school and community partners, and others defined by the provider, are involved in program evaluation, improvement, and identification of models of excellence.

Quality and Strategic Evaluation

5.1 The provider's quality assurance system is comprised of multiple measures that can monitor advanced program candidate progress, advanced completer achievements, and provider operational effectiveness. Evidence demonstrates that the provider satisfies all CAEP standards.

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.

Continuous Improvement

5.3 The provider regularly and systematically assesses performance against its goals and relevant standards, tracks results over time, tests innovations and the effects of selection criteria on subsequent progress and completion, and uses results to improve program elements and processes.

5.4 Measures of advanced program completer impact on the P-12 learning environment, including available outcome data on P-12 student growth, are summarized, externally benchmarked, analyzed, shared widely, and acted upon in decision-making related to programs, resource allocation, and future direction.

5.5 The provider assures that appropriate stakeholders, including alumni, employers, practitioners, school and community partners, and others defined by the provider, are involved in program evaluation, improvement, and identification of models of excellence.

Guidelines for Developing Assessment Instruments

From CAEP Evidence Manual, Section 6, p. 22: “Evidence Created and Administered by EPPs”

1. HOW THE ASSESSMENTS ARE USED

1. Is the point in the curriculum at which the assessment is administered clear (e.g. first year, last year, etc.)?
 - a. At entry, exit, mid-point, etc.?
 - b. While the emphasis should be on exit, are there examples of assessments or assignments at other points?
 - c. Are the curricular points an identified part of a clear developmental sequence?

NOTE: This information would be part of the documentation that the assessments are relevant.

2. HOW THE INSTRUMENTS ARE CONSTRUCTED

1. Are assessments aligned with CAEP Standards and not treated as a substitute for Standards? If so, then:
 - a. the same or consistent categories of content appear in the assessment that are in the Standards;
 - b. the assessments are congruent with the complexity, cognitive demands, and skill requirements described in the Standards; and that
 - c. the level of respondent effort required, or the difficulty or degree of challenge of the assessments, is consistent with Standards and reasonable for candidates who are ready to teach or to take on other professional educator responsibilities.

NOTE: Information on these aspects of assessments can be used by the provider to demonstrate construct or content validity and relevance.

3. HOW THE INSTRUMENTS ARE SCORED

1. Is there a clear basis for judging the adequacy of candidate work?
 - a. A rubric or scoring guide is supplied.
 - b. Multiple raters or scorers are used.
 - c. There is evidence that the assignment measures what it purports to measure

(NOTE: this information would be part of the evidence for construct validity or content validity and relevance) and that results are consistent across raters and over time (NOTE: this would be evidence of reliability).

- d. If good performance on one attribute can make up for poor performance on another, the EPP self-study explains the implications in terms of readiness to teach.
- e. If weights are used, they are explained or justified.
2. What do the performance levels represent?
 - a. There are three, four or five distinct levels, and they are clearly distinguishable from one another.
*NOTE: Our EPP recommendation is **FOUR**; a well-developed **THREE** level rubric may be okay.*
 - b. Levels are constructed in parallel with one another in terms of the attributes and descriptors used.
 - c. For each level of performance, attributes are described that are related to actual classroom performance; attributes are not simply mechanical counts of particular attributes.
 - d. Levels represent a developmental sequence in which each successive level is qualitatively different from the prior level.
 - e. Headings clearly describe which levels are acceptable and which are not acceptable.
 - f. It is clear which level represents exit proficiency (ready to practice).
 - g. A “no data” or “unobserved” category is included.

NOTE: Information in this category would help documents that the evidence is actionable—it is in forms directly related to the preparation program and can be used for program improvement and for feedback to the candidate.

3. Are the levels described in language that is readily understandable?
 - a. The levels should communicate to broad audiences including educators, stakeholders, and school partners.
 - b. Any special terms used are clearly defined.
4. Is there evidence of efforts to achieve consistency in scoring?
 - a. Multiple scorers are used.
 - b. Consistent training of reviewers is present.
 - c. Evidence of consistency such as inter-rater reliability is supplied.

NOTE: This information can be used by the provider to document reliability of the assessment.

4. HOW THE DATA ARE REPORTED

1. Are data reported?
 - a. Data are needed to show that the assessment is actually in use.
 - b. Data distributions (e.g. across rubric levels, disaggregated by area of specialty/ licensure preparation and by demographic groups) are reported and interpreted.
 - c. The EPP uses the data and its interpretation to suggest changes in the preparation program.
 - d. All candidates who completed the assessment are included or the cases included constitute a representative sample.

NOTE: this information would be appropriate for the providers to use in demonstrating that the data are representative.

2. How are results aggregated for reporting?
 - a. Scores are reported in terms of a percentage distribution of candidates scoring at each level or a mean with a range and not just a single central tendency (e.g. mean).
3. Are there comparisons?
 - a. EPP explains how it determines that an answer is “good enough”.
 - b. Comparisons should be criterion based.
 - c. The EPP describes other kinds of comparisons that are used (e.g. fixed standard or target, normative, improvement over time, comparison with peers in a state or region or nationally).

NOTE: The information from reporting is linked with the actionability principle since it determines how closely the information aligns with particular preparation programs or experiences and with groups of candidates.

5. INFORMING THE TEST TAKERS

1. Is there a mechanism for supplying feedback?
 - a. To candidates.
 - b. To the EPP for purposes of continuous improvement.
2. Are candidates given information about the bases on which they will be scored/ judged?

NOTE: This information can be used by the provider as part of their documentation that assessments are fair.

RUBRICS FOR EVALUATION OF EPP INSTRUMENTS USED AS ACCREDITATION EVIDENCE

**For use with: assessments, assignments, observation protocols, scoring guides and surveys created by EPPs
For use by: CAEP reviewers in Optional Early Instrument Evaluation
and CAEP Visitor Teams in review of self-studies**

EXCERPT from the CAEP HANDBOOK on “Optional Early Instruments Evaluation”

Early in the accreditation process, providers can elect to submit to CAEP the generic assessments, surveys, and scoring guides that they expect to use to demonstrate that they meet CAEP standards. . . The purpose of this review is to provide EPP’s with formative feedback on how to strengthen assessments, with the ultimate goal of generating better information on its candidates and continuously improving its programs.

Providers submit for review only the provider-created assessments used across all specialty/ license areas. This evaluation creates opportunities for providers to modify those instruments and begin to gather data with them that will be reported in the self-study and reviewed during the CAEP visit. This feature is a part of CAEP’s specialty/ license area review under Standard 1. . .

The array of categories contained in this Assessment Rubric is purposefully aligned with the CAEP Handbook description on the contents of submissions for the optional Early Instrument Evaluation. Submissions are to include (1) instruments (assessments, assignments, work samples, observations, surveys, etc.), (2) scoring guides, and (3) information about the standards that are informed by these instruments: (a) which items provide evidence for individual CAEP standards; (b) how the quality of the instrument/ evidence has been, or will be, determined; (c) the criteria for success measured for scoring guides and survey data, and (d) how the instruments were developed. The ten rubrics are constructed as reviewer guides for all parts of the Early Instruments Evaluation submission. They are grouped under five headings:

- A. Rubrics for EPP submissions on Instrument purpose, development and respondent information (categories 1-3);
- B. Rubrics for assessments, assignments and observation protocols (categories 4 and 5);
- C. Rubrics for scoring guides (categories 6 and 7);
- D. Rubrics for surveys (category 8); and
- E. Rubrics for validity and reliability (categories 9 and 10).

And a reminder for EPPs and reviewers: No single instrument can address all the content, complexity and difficulty contained in standards. Instead the cumulative assessments administered by the EPP should represent the range of standards. Providers should take this into account when they excerpt information from instrument results to document aspects of standards, and then, again, when they demonstrate for Standard 5 that their assessments are cumulative and coherent.

Assessment Rubrics

See the CAEP Evidence Guide section 5, “Validity and Other Principles of Good Evidence”, pp. 16-21, for additional definitions and descriptions. See section 6, pp. 22-26 for criteria to guide creation and use of assessments, scoring guides and surveys.

Category	Level 1	Level 2	Level 3	Level 4	Reviewer Comments
Rubric number, category and description; reference to evidence principles addressed	<i>Does not meet the minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	<i>Approaches minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	<i>Meets minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	<i>Demonstrates target criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines at a high level of performance.</i>	
A. RUBRICS FOR EPP SUBMISSIONS ON INSTRUMENT PURPOSE, DEVELOPMENT AND RESPONDENT INFORMATION					
1. INSTRUMENT PURPOSE AND USE: Administration of the instrument in the program, its purpose, and standards addressed (informs relevance, content validity)	<ul style="list-style-type: none"> Use of the instrument during preparation is generally described or ambiguous The purpose of the instrument and its use in candidate monitoring or decisions on progression are generally described Specific standards addressed by the instrument are not provided 	<ul style="list-style-type: none"> Use of the instrument during preparation is generally described but not in terms of the sequence of candidate progression The purpose of the instrument is described only in general terms without reference to particular candidate decisions to be made Specific standards addressed by the instrument are not clearly identified 	<ul style="list-style-type: none"> The point or points when the instrument is administered during the preparation program are explicit The purpose of the instrument and its use in candidate monitoring or decisions on progression are specified The CAEP, InTASC or State standards that the instrument will inform are explicit 	<ul style="list-style-type: none"> The point when the instrument is administered during the preparation program are explicit Candidate progression is monitored and the information used for mentoring The purpose of the instrument and its use in candidate monitoring or decisions on progression are specified and decisions are consequential The CAEP, InTASC or State standards that the instrument will inform are explicit 	
2. INSTRUMENT DEVELOPMENT: How the instrument was developed (informs	<ul style="list-style-type: none"> EPP provides limited description of instrument’s development 	<ul style="list-style-type: none"> EPP provides a description of the instrument’s development 	<ul style="list-style-type: none"> EPP provides a detailed description of the instrument’s development 	<ul style="list-style-type: none"> EPP provides a description of the instrument’s development indicating 	

Assessment Rubrics

Category Rubric number, category and description; reference to evidence principles addressed	Level 1 <i>Does not meet the minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 2 <i>Approaches minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 3 <i>Meets minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 4 <i>Demonstrates target criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines at a high level of performance.</i>	Reviewer Comments
(relevance)	<ul style="list-style-type: none"> • No evidence is provided that the instrument is integrated with aspects of preparation curriculum • EPP has provide no information to indicate faculty input or concurrence 	<ul style="list-style-type: none"> • Limited evidence to indicate that the instrument is integrated with preparation curriculum • Evidence indicates that instrument development was not conducted with wide faculty input and concurrence 	<ul style="list-style-type: none"> • Instrument development is integrated with preparation curriculum • Instrument development engaged relevant preparation provider and clinical faculty 	stages for piloting and refinements <ul style="list-style-type: none"> • Instrument development is integrated with preparation curriculum and stages of candidate progression • Instrument development engaged relevant preparation provider and clinical faculty at multiple stages 	
3. INFORMATION FOR RESPONDENTS: information given to respondent before and at the administration of the instrument (informs fairness and reliability)	<ul style="list-style-type: none"> • EPP provides little or no general information to respondents about the purpose of the results from the instrument • Instructions provided to respondents are incomplete and/ or ambiguous • Information is not provided about how respondents' work will be judged 	<ul style="list-style-type: none"> • EPP provides general information to the respondents about the purpose of the results from the instrument • Instructions provided to respondents are incomplete and/ or ambiguous • Sketchy information is provided about how respondents' work will be judged 	<ul style="list-style-type: none"> • 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 basis for judgment (criterion for success, or what is "good enough") is made explicit for respondents 	<ul style="list-style-type: none"> • The respondents for the instrument are given a description of its purpose • Respondents are informed how the instrument results are used in reaching conclusions about their status and/ or progression in the preparation program • Instructions provided to respondents about what they are expected to do are informative and 	

Assessment Rubrics

Category	Level 1	Level 2	Level 3	Level 4	Reviewer Comments
Rubric number, category and description; reference to evidence principles addressed	<i>Does not meet the minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	<i>Approaches minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	<i>Meets minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	<i>Demonstrates target criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines at a high level of performance.</i>	
				unambiguous <ul style="list-style-type: none"> The basis for judgment (criterion for success or what is “good enough”) is made explicit for respondents 	
B. RUBRICS FOR ASSESSMENTS, ASSIGNMENTS, AND OBSERVATION PROTOCOLS					
B.4.1 Alignment with standards					
4. ASSESSMENTS and ASSIGNMENTS: Alignment with standard (informs content and construct validity and relevance) [Repeating a note from the introduction: No single instrument can address all the content, complexity and difficulty contained in the standards. Instead the cumulative assessments administered by the EPP should represent the range of standards.]	The assessment items, or the assignment tasks, are: <ul style="list-style-type: none"> Only occasionally consistent with the content of the standards being informed; Represent only few of the complexity or cognitive demands found in the standards, and Fail to reflect the degree of difficulty or level of effort described in the standards. 	The assessment items, or the assignment tasks, are: <ul style="list-style-type: none"> usually consistent with the content of the standards being informed; represent most of the range of complexity or cognitive demands found in the standards, and partially reflect the degree of difficulty or level of effort described in the standards. 	The assessment items, or the assignment tasks, are: <ul style="list-style-type: none"> consistent with the content of the standards being informed; represent the complexity or cognitive demands found in the standards, and reflect the degree of difficulty or level of effort described in the standards. 	The assessment items, or the assignment tasks, are: <ul style="list-style-type: none"> consistent with the content of the standards being informed; represent the complexity or cognitive demands found in the standards, and reflect the degree of difficulty or level of effort described in the standards. 	
	B.4.2 Representation of criteria, especially for higher level functioning				
	<ul style="list-style-type: none"> Alignment criteria are demonstrated rarely or not at all (less than 25%). 	<ul style="list-style-type: none"> Alignment criteria are demonstrated only inconsistently (25% to 49%) 	<ul style="list-style-type: none"> Alignment criteria are consistently demonstrated (50% to 75%) 	<ul style="list-style-type: none"> Alignment criteria are consistently demonstrated (75% or more) 	

Assessment Rubrics

Category Rubric number, category and description; reference to evidence principles addressed	Level 1 <i>Does not meet the minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 2 <i>Approaches minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 3 <i>Meets minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 4 <i>Demonstrates target criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines at a high level of performance.</i>	Reviewer Comments
	<ul style="list-style-type: none"> Assessments and assignments include few items that reflect the complexity, cognitive demands and difficulty of the standard/ components. Standard/ components that require higher levels of intellectual behavior (e.g., create, evaluate, analysis, & apply) are not prevalent in the assessment/ assignment, which instead represents identify, remember, and understand. For example, when a standard requires candidates' students to "demonstrate" problem solving, the item on the assessment has candidates requiring students only to complete worksheets or identify specific content. 	<ul style="list-style-type: none"> Assessments and assignments include less than a majority of items that are congruent with standard/ components that require higher levels of intellectual behavior (e.g., create, evaluate, analysis, & apply) and more items representative of identification, remembering and understanding skills. For example, when a standard requires candidates' students to "demonstrate" problem solving, the item on the assessment has candidates requiring students only to complete worksheets or identify specific content. 	<ul style="list-style-type: none"> Assessments and assignments include items congruent with standard/ components that require higher levels of intellectual behavior (e.g., create, evaluate, analysis, & apply). For example, when a standard requires candidates' students to "demonstrate" problem solving, then the assessment item is specific to students' application of knowledge to solve problems. 	<ul style="list-style-type: none"> Assessments and assignments include items congruent with the complexity, cognitive demands, and/or skills required and are linked to challenging and innovative learning experiences. For example, when a standard requires candidates' students to "demonstrate" problem solving, then candidates ask students to "use" or "apply" content knowledge in a project-based learning experience across more than one discipline. 	
5. OBSERVATION PROTOCOLS:	B.5.1 Alignment with standards				
	<ul style="list-style-type: none"> Reviewer protocols 	<ul style="list-style-type: none"> Reviewer protocols 	<ul style="list-style-type: none"> Reviewer protocols 	<ul style="list-style-type: none"> Reviewer protocols 	

Assessment Rubrics

Category	Level 1	Level 2	Level 3	Level 4	Reviewer Comments
Rubric number, category and description; reference to evidence principles addressed	<i>Does not meet the minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	<i>Approaches minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	<i>Meets minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	<i>Demonstrates target criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines at a high level of performance.</i>	
Alignment with standards and good data practices (informs relevancy) and information for the candidate (informs fairness)	contain evaluation categories that are not shown to be in alignment with CAEP, InTASC and/or State standards	contain evaluation categories only generally aligned with CAEP, InTASC and/or State standards	contain evaluation categories clearly aligned with CAEP, InTASC and/or State standards	contain evaluation categories clearly aligned with CAEP, InTASC and/or State standards	
B.5.2 Clarity and significance of the observation categories					
[NOTE: Rubrics in this row address the construct of the observer’s protocol. See “Scoring”, items 6 and 7, for rubrics on the levels of judgment and “Reliability”, item 10, on training of observers.]	<ul style="list-style-type: none"> Evaluation categories are not described or described only in ambiguous language Half or more of the evaluation categories require observers to judge attributes of candidate proficiencies that are of less importance in the standards 	<ul style="list-style-type: none"> Evaluation categories are described but sometimes in ambiguous language Some evaluation categories (25% or more of total score) require observers to judge attributes of candidate proficiencies that are of clearly less importance in the standards 	<ul style="list-style-type: none"> Evaluation categories unambiguously describe the proficiencies to be evaluated Most evaluation categories (80% of the total score) require observers to judge consequential attributes of candidate proficiencies in the standards 	<ul style="list-style-type: none"> Evaluation categories unambiguously describe the proficiencies to be evaluated Almost all evaluation categories (95% of the total score) require observers to judge consequential attributes of candidate proficiencies in the standards 	
C. RUBRICS FOR SCORING GUIDES					
6. SCORING LEVELS: Candidate proficiency levels are clearly distinguishable (informs reliability, and also evidence principle of “actionability” in decisions about	C. 6.1 Rating scales				
	<ul style="list-style-type: none"> Rating scales are used in lieu of rubrics. These rating scales use a single definition for each level that is applied to all items on the assessment. For 	<ul style="list-style-type: none"> Vague, general terms are used to differentiate levels. These terms are open to multiple interpretations, which limits the reliability of 	<ul style="list-style-type: none"> Levels are qualitatively defined using specific criteria aligned with key attributes identified in the item. Levels represent a developmental 	<ul style="list-style-type: none"> Levels are qualitatively defined using specific criteria aligned with key attributes identified in the item. By qualitatively defining performance at each 	

Assessment Rubrics

Category Rubric number, category and description; reference to evidence principles addressed	Level 1 <i>Does not meet the minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 2 <i>Approaches minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 3 <i>Meets minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 4 <i>Demonstrates target criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines at a high level of performance.</i>	Reviewer Comments
programs and candidates) and reviewers are trained (informs reliability)	example, level 1 = significantly below expectation; level 2 = below expectation; level 3 = meets expectation; level 4 = significantly above expectation. Levels do not represent a qualitative difference from the prior level. Rating scales provided no feedback to candidates specific to their performance on each item.	the assessment and provides limited feedback to candidates. For example, levels are differentiated by: level 1 – “no understanding”; level 2 – “limited understanding”; level 3 – “understanding”; level 4 – “complete understanding.” The criteria remain the same at each level of the rubric with qualitative differentiation defined by vague terms that provide limited feedback and guidance to candidates.	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.	level, candidates are provided with descriptive feedback on their performance and consistency across raters is increased. <ul style="list-style-type: none"> • Criteria for each attribute in the item are identified. • Multiple raters are trained and used 	
	C.6.2 Training scorers				
7. SCORING ATTRIBUTES:	<ul style="list-style-type: none"> • No evidence on training of raters or scorers or on inter-rater reliability 	<ul style="list-style-type: none"> • Only informal evidence of attempts to ensure inter-rater reliability in scoring 	<ul style="list-style-type: none"> • Multiple raters or scorers are trained and used 	<ul style="list-style-type: none"> • Results are monitored over time and compared with standardized scoring 	
	<ul style="list-style-type: none"> • No clear basis for judging candidate work 	<ul style="list-style-type: none"> • The basis for judging candidate work is 	<ul style="list-style-type: none"> • The basis for judging candidate work is well 	<ul style="list-style-type: none"> • The basis for judging candidate work is well 	

Assessment Rubrics

Category Rubric number, category and description; reference to evidence principles addressed	Level 1 <i>Does not meet the minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 2 <i>Approaches minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 3 <i>Meets minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 4 <i>Demonstrates target criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines at a high level of performance.</i>	Reviewer Comments
Proficiency levels are stated in performance or observable behavior terms (informs principle of “actionability” in program or candidate decisions)	is defined. <ul style="list-style-type: none"> Does not provide actionable feedback to candidates Performance attributes are not defined, but simply repeated from the standard/ component. 	vague and ill-defined. <ul style="list-style-type: none"> Does not provide actionable feedback to candidates Performance attributes are defined using vague terms that are not actionable, performance based, or in observable behavior terms. Items use such terms as “understand” or “learns”. 	defined <ul style="list-style-type: none"> 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. 	defined <ul style="list-style-type: none"> Feedback is provided to candidates is actionable Performance attributes are defined in actionable, performance based or observable behavior terms. Higher level action verbs from Bloom’s taxonomy are used throughout assessments such as “application of knowledge” or “analysis”. If less actionable term is used such as “engaged”, criteria are provided to define the use of the term in the context of the item. 	
D. RUBRICS FOR SURVEYS					
8. SURVEY CONSTRUCTION AND INSTRUCTIONS: Instruments are constructed to follow sound survey research practice and	D.8.1 Survey item construction				
	<ul style="list-style-type: none"> Individual items or questions do not use clear language and may include items with more than one subject. Items are usually stated 	<ul style="list-style-type: none"> Individual items or questions usually have a single subject but are sometimes ambiguous Items are sometimes stated in terms of 	<ul style="list-style-type: none"> Individual items or questions are simple and direct; Questions have a single subject; language is unambiguous. 	<ul style="list-style-type: none"> Individual items or questions are simple and direct; Questions have a single subject; language is unambiguous. 	

Assessment Rubrics

Category Rubric number, category and description; reference to evidence principles addressed	Level 1 <i>Does not meet the minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 2 <i>Approaches minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 3 <i>Meets minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 4 <i>Demonstrates target criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines at a high level of performance.</i>	Reviewer Comments
completers are given information about the survey's purpose (informs relevance)	in terms of opinions, rather than as behaviors or practices	behaviors or practices	<ul style="list-style-type: none"> Items are stated in terms of behaviors or practices instead of opinions, whenever possible 	<ul style="list-style-type: none"> Items are stated in terms of behaviors or practices instead of opinions, whenever possible Scoring is anchored in performance or behavior demonstrably related to teaching practice Questions follow a parallel structure. Leading questions are avoided. 	
D.8.2 Disposition surveys information for respondents					
	<ul style="list-style-type: none"> Surveys of dispositions provide no explanations of the purpose of the survey. 	<ul style="list-style-type: none"> Surveys of dispositions fail to specify how the survey information is related to effective teaching. 	<ul style="list-style-type: none"> Surveys of dispositions make clear to respondents how the survey is related to effective teaching 	<ul style="list-style-type: none"> Surveys of dispositions make clear to respondents how the survey is related to effective teaching and impact on P-12 student learning. 	
E. RUBRICS FOR VALIDITY AND RELIABILITY					
9. INSTRUMENT VALIDITY: Degree to which an assessment measures what it purports to measure and how the results will be interpreted	<ul style="list-style-type: none"> No description or plan is provided for establishing validity for the instrument The instrument was not piloted prior to administration 	<ul style="list-style-type: none"> A description or plan is provided that is non-specific or fails to provide enough information for reviewers to determine whether validity is 	<ul style="list-style-type: none"> A description or plan is provided that details steps the EPP has taken or is taking to ensure the validity of the assessment The plan details the 	<ul style="list-style-type: none"> A description or plan is provided that details steps the EPP has taken or is taking to ensure the validity of the assessment The plan details the 	

Assessment Rubrics

Category Rubric number, category and description; reference to evidence principles addressed	Level 1 <i>Does not meet the minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 2 <i>Approaches minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 3 <i>Meets minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 4 <i>Demonstrates target criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines at a high level of performance.</i>	Reviewer Comments
(informs principle of validity)		under investigation or has been established. The instrument was not piloted prior to administration <ul style="list-style-type: none"> • Description or plan not specific, or described steps do not meet accepted research standards • Validity is determined by an internal review by one or two stakeholders. For example, the EPP notes that validity was established since the assessment was reviewed by the dean and associate dean. 	types of validity that are under investigation or have been established (e.g., construct, content, concurrent, predictive, etc.) <ul style="list-style-type: none"> • The instrument was developed drawing on research about content and format • The instrument was piloted prior to administration • The EPP details its plans for analyzing and interpreting results from the instrument. • The described steps generally meet accepted research standards for establishing the validity of an assessment. 	types of validity that are under investigation or have been established (e.g., construct, content, concurrent, predictive, etc.) <ul style="list-style-type: none"> • The instrument was developed drawing on research about content and format • The instrument was piloted prior to administration • The EPP details its plans for analyzing and interpreting results from the instrument. • The described steps meet accepted research standards for establishing the validity of an assessment. • A validity coefficient is reported. 	
10. INSTRUMENT RELIABILITY: Degree to which an assessment produces stable and consistent	<ul style="list-style-type: none"> • No description or plan is provided for establishing reliability for the assessment. • No evidence that 	<ul style="list-style-type: none"> • A description or plan is provided that is non-specific or fails to provide enough information to 	<ul style="list-style-type: none"> • A description or plan is provided that details the type of reliability that is being investigated or has 	<ul style="list-style-type: none"> • A description or plan is provided that details the type of reliability that is being investigated or has been 	

Assessment Rubrics

Category Rubric number, category and description; reference to evidence principles addressed	Level 1 <i>Does not meet the minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 2 <i>Approaches minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 3 <i>Meets minimum criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines.</i>	Level 4 <i>Demonstrates target criteria necessary to support a CAEP evaluation concluding that self-study data are likely to meet CAEP standards and evidence guidelines at a high level of performance.</i>	Reviewer Comments
results. Answers the question – “Can the evidence be corroborated?”	scorers are trained	determine if reliability is being investigated or has been established. <ul style="list-style-type: none"> • The specific type of reliability is not identified (e.g., test-retest, parallel forms, inter-rater, internal consistency, etc.) • Little or no evidence that scorers are trained • The described steps are informal, and fall short of research standards. 	been established (e.g., test-retest, parallel forms, inter-rater, internal consistency, etc.) and the steps the EPP took to ensure the reliability of the assessment. <ul style="list-style-type: none"> • Training of scorers and checking on inter-rater reliability are documented • The described steps meet accepted research standards for establishing reliability 	established (e.g., test-retest, parallel forms, inter-rater, internal consistency, etc.) and the steps the EPP took to ensure the reliability of the assessment. <ul style="list-style-type: none"> • Training of scorers and checking on inter-rater reliability are documented • The described steps meet accepted research standards for establishing reliability • A reliability coefficient is reported. 	