

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 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
				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					
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.]	B.4.1 Alignment with 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	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>		
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					
	<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 		
7. SCORING ATTRIBUTES:	<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. 	