



Council for the
Accreditation of
Educator Preparation

Quality of Evidence

Christine Carrino Gorowara, CAEP
Vice President for the IB Pathway and CAEP
Evidence

Session Overview

- This session will focus on
 - Qualities of evidence
 - Benefits of Anchor Measures in establishing a body of good evidence from which to draw
 - Analyzing evidence quality to examine the strength and appropriateness evidence sources

“Evidence is not something that an EPP ‘does for the accreditor.’ It is not a ‘compliance’ mechanism. The data are not an end in themselves or ‘the answer’ for accreditation. Instead, data are the basis to begin a conversation.”

CAEP Evidence Guide, http://caepnet.files.wordpress.com/2014/02/caep_evidence_guide1.pdf

Standard 5, Component 5.2

The provider's quality assurance system relies on **relevant, verifiable, representative, cumulative** and **actionable** measures, and produces empirical evidence that interpretations of data are **valid** and **consistent**.

Valid and Consistent

- Validity is about the appropriateness of the results of an assessment for a specific purpose
- Validity depends in part upon the reliability of the measure
- Validity is a property of the assessment results and interpretation, not of the instrument
- Consistency, or reliability, can be measured across time, across/within raters, within assessment instrument
- High consistency or reliability does not compensate for low validity

Relevant

- Relevance is related to validity
- Relevant measures are demonstrably related to a question of importance that is being investigated
- There must be a clear and explicable link between what a particular measure is established to gauge and the substantive content of the Standard under which it is listed

Verifiable

- The validity of any measure advanced should be able to be independently verified.
- Verifiability requires transparency and full documentation, as well as ability for the assessment to be replicated.

Representative

- Any measure put forward should be typical of an underlying situation or condition, not an isolated case.
- Data should be provided that shows that the case or sample is sufficiently similar to the population.
- Evidence of representativeness must be subject to verification by a third party.

Note: sometimes it is easier to include the entire population than to select a representative sample, so sampling should be considered when assessing the entire population is challenging.

Cumulative

- Measures gain credibility as additional sources or methods for generating them are employed.
- The resulting triangulation helps guard against the inevitable flaws associated with any one approach.
- The entire set of measures used under a given Standard should be mutually reinforcing.

Actionable

- Good measures provide programs with specific guidance for action and improvement.
- Any measure should be able to be disaggregated to reveal underlying patterns of strength and weakness or to uncover populations who could be served more effectively.
- The measures provided should be reflectively analyzed and interpreted to reveal specific implications for the program.

Anchor Measures

- In an ideal data system there would be a few common measures
 - Serve as anchors—informative, but also help to calibrate other measures
- One current example—licensure tests
 - e.g., Praxis English Language, literature and composition content knowledge
 - 24296 candidates took the test in 2012-2013; 7 states require the test
 - Median score 176; 25th to 75th percentile range 164-187
 - Construct: literature and understanding text = 55%; language and linguistics = 15%; composition and rhetoric = 30%
 - How does the construct align with your curriculum?
 - Is this an appropriate measure of how your EPP prepares teachers?

Anchor Measures, continued

- Similar measures for Pearson
- Examples of other measures that could become “anchors” for preparation comparisons, benchmarks, continuous improvement and accreditation
 - edTPA or Praxis Performance Assessments for Teachers
 - New developments in licensure tests
 - CAEP’s “8 annual measures” of program impact and program outcome

Evidence for making the case for accreditation

- The responsibility of making the case for accreditation lies with the EPP. It is not enough to present CAEP with sources of evidence listed in its guidance documents.
- Rather, the EPP should choose a set of evidence that makes a compelling case for each Standard.
- Accordingly, an institution is welcome to employ different practices from those described in CAEP guidance documents; in that case, the institution is responsible for showing that it has addressed the intent of each Standard in an equally effective way.

Standards and components

- Accreditation depends on demonstrating that the EPP has met each Standard. Evidence should be consistent with the Standard's holistic and overarching expectation.
- While each component should be addressed in some way, different EPPs may give different weights to individual components.

Tips Regarding Evidence

- Consider relevance first, then validity/reliability (trustworthiness/consistency).
- Even high-quality evidence will have relative strengths and weaknesses. Balance evidence that is weak in one area with evidence that is strong in that area.
- Be sure that all sources of evidence are of sufficient quality, though: increasing the number of irrelevant, invalid, or unreliable pieces of evidence does not strengthen the EPP's case.

Follow-Up

Engaged feedback is vital to CAEP. You will have an opportunity to complete a survey at the end of the conference. Surveys will be sent via email on Friday, April 10. We encourage your participation. Thank you.

Further questions:

christine.gorowara@caepnet.org