RETHINKING THE INTERN EVALUATION TO BETTER PREDICT IMPACT: ONE INSTITUTION’S SELECTED IMPROVEMENT PLAN

Elayne Colón and Tom Dana
University of Florida

CAEPCon, Fall 2017
Washington, DC
Institutional Overview and Context

- NCATE/CAEP accredited since 1954 – NCATE review in 2010, CAEP review in 2017
- 14 State-approved teacher education programs
- Distinct national and state reviews (not a “SPA State”); no partnership agreement
- ~400 candidates enrolled in teacher education programs in 2016-17
- 173 completers of state-approved teacher education programs in 2016-17
CAEP Accreditation Review Timeline

- Summer 2015: Early Assessment Review by CAEP
- August 2016: Self-Study & Selected Improvement Plan Submitted
- Fall 2016: Off-site Review
- January 2017: Off-site Report Received
- April 1-4, 2017: On-site Review
- Fall 2017: Final Accreditation Decision
The **Selected Improvement Pathway** asks the provider to select a standard or standards and/or components across standards and develop an improvement plan that addresses them and uses evidence from the self-study to demonstrate improvement.

~ http://caepnet.org/accreditation/caep-accreditation/caep-accreditation-resources/selected-improvement
Selected Improvement Plan

Goals –
1) Improve the reliability of the Intern Evaluation (IE) instrument through revisions to instrument and revised training materials for supervisors
2) Determine the predictive validity of the revised IE
Rationale and Supporting Research for SI Plan

- Internship (i.e., Student Teaching) cited as one of the most influential aspects of teacher preparation (e.g., National Research Council, 2010)
- UF IE serves as the final, high-stakes assessment of candidates’ performance
- Important to understand relationship between university and field-based supervisor ratings of interns and implications for future teaching performance
- Impact: Teacher effectiveness is the most important school-based factor associated with student achievement (Goldhaber, Krieg, & Theobald, 2016)
- Confidence in IE ratings (i.e., validity) supports use of data for program evaluation/improvement efforts
CAEP Standards Aligned to Selected Improvement Plan

- **Standard 2.3 (Clinical Experiences)** – The provider utilizes multiple performance-based assessments that demonstrate candidates’ development of the knowledge, skills, and professional disposition associated with impact on learning and development of all P-12 students.

- **Standard 5.2 (Quality and Strategic Evaluation)**: The provider’s quality assurance system relies on relevant, verifiable, representative, cumulative and actionable measures, and produces empirical evidence that interpretations are valid and consistent.

- **Standard 5.3 (Continuous Improvement)** – The provider documents that it regularly and systematically tests innovations and uses results to improve program elements.
Early Assessment Review

“The utility of educator preparation provider (EPP) data used for continuous improvement of candidates and providers, as well as evidence in the accreditation process, is important to CAEP. Quality assessments are critical to these purposes. Therefore, we strongly encourage EPPs to conduct reviews of their assessments and to employ experts – either internally or within the education field – as needed.”

http://caepnet.org/accreditation/caep-accreditation/caep-accreditation-resources
CAEP Evaluation Framework for EPP-Created Assessments

- For use with: Educator preparation provider (EPP)-created assessments, including subject and pedagogical content tests, observations, projects, assignments, and surveys.
- For use by: EPPs to evaluate their own assessments and by CAEP site teams to review evidence in self-study submissions.

http://caepnet.org/accreditation/caep-accreditation/caep-accreditation-resources
CAEP Evaluation Framework for EPP - Created Assessments

3. SCORING (informs reliability and actionability)

see p. 2 handout
3. SCORING (informs reliability and actionability)

a. The basis for judging candidate performance is well defined.
b. Each Proficiency Level Descriptor (PLD) is qualitatively defined by specific criteria aligned with indicators.
c. PLDs represent a developmental sequence from level to level (to provide raters with explicit guidelines for evaluating candidate performance and for providing candidates with explicit feedback on their performance).
d. Feedback provided to candidates is actionable—it is directly related to the preparation program and can be used for program improvement as well as for feedback to the candidate.
e. Proficiency level attributes are defined in actionable, performance-based, or observable behavior terms. [NOTE: 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 category or indicator.]

a. Higher level actions from Bloom’s or other, taxonomies are used in PLDs such as “analyzes” or “evaluates.”
FEAP 6: Professional Responsibility & Conduct

6a. The effective educator adheres to the Code of Ethics and the Principles of Professional Conduct of the Education Profession of Florida

Part 2: UF Teacher Professionalism Indicators

As a professional, the effective educator:

2.1 Reflects on the extent to which learning goals were met and how instruction can be changed to facilitate learning

2.2 Demonstrates a sense of efficacy

2.3 Builds professional relationships with colleagues to share teaching insights and to coordinate learning activities for students

2.4 Demonstrates initiative and self-reliance

2.5 Demonstrates enthusiasm for teaching

2.6 Demonstrates responsiveness to supervision (ability to accept constructive criticism and incorporate suggestions into teaching performance)

2.7 Demonstrates responsibility for maintaining accurate student records and other important information

2.8 Is punctual

2.9 Presents a professional appearance in dress, grooming, attitude, and demeanor

OVERALL DOMAIN RATING:

Comments:

UF College of Education

Student Teaching Evaluation

Candidate Name (Last, First):

Submitted by (Check One): Directing Teacher ☐ UF Supervisor ☐ Other: ____________________________

Name of Educator: ____________________________ Placement Location and Grade Level(s): ________________

This form is to serve as a permanent record of the summative evaluation of a teacher candidate’s performance and demonstration of competence on the Florida Educator Accomplished Practices as part of the culminating field experience. Please rate the candidate’s performance on each indicator as well as an overall domain rating, when applicable.

Unsatisfactory Developing Accomplished Exceptional

The candidate demonstrates little knowledge of this practice. The candidate usually demonstrates this practice. The candidate is independent in routine situations with minimal to no supervision required.

Part 1: Florida Educator Accomplished Practices (FEAP)

FEAP1: Instructional Design and Planning

Applying concepts from human development and learning theories, the effective educator consistently:

1a. Aligns instruction with state-adopted standards at the appropriate level of rigor

1b. Sequences lessons and concepts to ensure coherence and required prior knowledge

1c. Designs instruction for students to achieve mastery

1d. Selects appropriate formative assessments to monitor learning

1e. Uses diagnostic student data to plan lessons

1f. Develops learning experiences that require students to demonstrate a variety of applicable skills and competencies.

OVERALL FEAP 1 RATING:

Signature: ____________________________ Date: __/__/____

Please return completed form to:

UNIVERSITY OF FLORIDA, COLLEGE OF EDUCATION, STUDENT SERVICES, G-415 NORMAN HALL
PO BOX 117042, GAINESVILLE, FL 32611-7042
Evolution of the Intern Evaluation Instrument

- February 2015: State Rule revised to require programs to use, “state-approved performance evaluation system that is aligned with a partnering school district(s)’ evidence-based framework” for final summative evaluation (Rule 6A-5.066, F.A.C.)
- Winter/Spring 2015: CAEP distributed assessment “rubric” recommending shift away from rating scales for EPP-created assessments
- Summer 2015: Submitted draft excerpt of Intern Evaluation (IE) aligning State Standards with Marzano and Danielson instructional frameworks to CAEP for Assessment Review
- 2015-16 Academic Year: Worked to draft, vet, and finalize detailed performance descriptions of four levels and each item on IE
- Fall 2016: First use of revised IE for all teacher education programs
Intern Evaluation

Candidate Name (Last, First): UFID: ____________________________
Name of Evaluator: ____________________________

Submitted by (Check One):  □ Directing ☐ Teacher □ UF Supervisor Other: ____________________________

Placement Location and Grade Level(s): ____________________________

This form is to serve as a permanent record of the summative evaluation of a teacher candidate’s performance and demonstration of competence on the Florida Educator Accomplished Practices as part of the culminating field experience. Please rate the candidate’s performance on each indicator, as well as an overall domain rating, when applicable.

Part 1: Florida Educator Accomplished Practices (FEAP)

FEAP 1: INSTRUCTIONAL DESIGN AND PLANNING

1a. Aligns instruction with state-adopted standards at the appropriate level of rigor:

<table>
<thead>
<tr>
<th>Unsatisfactory</th>
<th>Developing</th>
<th>Accomplished</th>
<th>Exceptional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructional plans lack alignment with the standards. Rigor of the learning goals and outcomes are not suitable for most students in the class.</td>
<td>Instructional plans demonstrate alignment with the standards. Learning goals/outcomes are based on a global assessment of student learning. Rigor is not suitable to account for the needs of diverse learners.</td>
<td>The teacher candidate demonstrates a thorough understanding of the concepts and skills s/he is teaching and what students are to learn (the Florida Standards). Learning goals/outcomes are aligned with the standards and student needs and represent rigorous and important learning in the discipline.</td>
<td>The teacher candidate demonstrates deep understanding of the concepts and skills s/he is teaching, what students are to learn, and interconnectedness of other disciplines and/or standards (the Florida Standards). Learning goals/outcomes are created aligned with the standards and student needs and represent rigorous and important learning in the discipline.</td>
</tr>
</tbody>
</table>

1b. Sequences lessons and concepts to ensure coherence and required prior knowledge:

<table>
<thead>
<tr>
<th>Unsatisfactory</th>
<th>Developing</th>
<th>Accomplished</th>
<th>Exceptional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lessons are disorganized, chaotic, and/or do not follow a recognizable progression. Prior knowledge is not accounted for in the structure of lesson(s).</td>
<td>Lessons have a recognizable structure that progress in a logical manner in accordance with the standards. Lessons adequately account for students' prior knowledge.</td>
<td>Lessons lead toward mastery of standards in a logical and coherent manner. Lessons access and expand upon prior knowledge with an understanding of prerequisite relationships within the content area, building upon each other in preparation for future learning.</td>
<td>Lessons lead toward mastery of the standards in a logical and coherent manner. Lessons access and expand upon prior knowledge with an understanding of prerequisite relationships within the content area and across the curriculum, building upon each other in preparation for future learning.</td>
</tr>
</tbody>
</table>
1f. Develops learning experiences that require students to demonstrate a variety of applicable skills and competencies. (Add Resources, Technology):

<table>
<thead>
<tr>
<th></th>
<th>Unsatisfactory</th>
<th>Developing</th>
<th>Accomplished</th>
<th>Exceptional</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning experiences are not designed to engage students in higher level skills that lead to mastery.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning experiences engage students in limited opportunities for higher level thinking with no differentiation for varying student needs. No plans exist for extension and application of skills and competencies that lead to mastery.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning experiences are designed to allow students opportunities for higher level thinking in a manner that is appropriate based on student needs. There are provisions for continued learning through expansion using a variety of skills and competencies that lead to mastery of academic standards.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning experiences engage students in higher-level cognitive activities, appropriately differentiated for individual learners. The lessons provides a variety of opportunities for students to apply and extend skills and understandings in related areas of the curriculum.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

OVERALL FEAP 1 RATING: Instructional Design and Planning

<table>
<thead>
<tr>
<th></th>
<th>Unsatisfactory</th>
<th>Developing</th>
<th>Accomplished</th>
<th>Exceptional</th>
</tr>
</thead>
</table>

FEAP 2: THE LEARNING ENVIRONMENT

2a. Organizes, allocates, and manages the resources of time, space, and attention:

<table>
<thead>
<tr>
<th></th>
<th>Unsatisfactory</th>
<th>Developing</th>
<th>Accomplished</th>
<th>Exceptional</th>
</tr>
</thead>
<tbody>
<tr>
<td>There is little evidence that the teacher candidate helps students understand routines and procedures that keep them focused and on task. Students are confused or unclear about simple procedures like lining up or how to transition from one activity to another. Time is not used efficiently. Physical classroom environment is poorly organized or unsafe, which inhibits developmentally appropriate instruction or movement. Teacher candidate is unaware of when students need attention.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The teacher candidate attempts to use routines to minimize loss of instructional time, but routines are not practiced or rehearsed enough and the classroom appears chaotic or poorly organized. Time is lost in transitions, and teacher candidate spends too much time reminding students of routines and procedures. Use of routines is uneven or inconsistently used throughout the day.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The teacher candidate reviews and monitors established classroom routines and procedures, rehearses routines and procedures with students, assesses students' understanding and proficiency in demonstrating routines and procedures, and employs routines and procedures, devoting more time to instruction and providing for a safe environment that promotes student learning.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical space is designed to be safe and well organized to minimize distractions and create a site conducive to learning. Students need little direction to perform routine tasks and procedures. The teacher candidate minimizes transition time by putting clear routines in place and maximizing every moment possible for learning. Resources and materials are clearly available and ready for all classroom activities to minimize time lost.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Selected Improvement Plan

Goals –
1) Improve the reliability of the Intern Evaluation (IE) instrument through revisions to instrument and revised training materials for supervisors
2) Determine the predictive validity of the revised IE
CAEP Evaluation Framework for EPP-Created Assessments

4. DATA RELIABILITY

5. DATA VALIDITY

see p. 3 handout
4. DATA RELIABILITY
   a. A description or plan is provided that details the type of reliability that is being investigated or has 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 data from the assessment.
   b. Training of scorers and checking on inter-rater agreement and reliability are documented.
   c. The described steps meet accepted research standards for establishing reliability.

   a. Raters are initially, formally calibrated to master criteria and are periodically formally checked to maintain calibration at levels meeting accepted research standards.
   b. A reliability coefficient is reported.

5. DATA VALIDITY
   a. A description or plan is provided that details steps the EPP has taken or is taking to ensure the validity of the assessment and its use.
   b. The plan details the types of validity that are under investigation or have been established (e.g., construct, content, concurrent, predictive, etc.) and how they were established.
   c. If the assessment is new or revised, a pilot was conducted.
   d. The EPP details its current process or plans for analyzing and interpreting results from the assessment.
   e. The described steps meet accepted research standards for establishing the validity of data from an assessment.

   a. Types of validity investigated go beyond content validity and move toward predictive validity.
   b. A validity coefficient is reported.
## SI Plan: Proposed Timeline

<table>
<thead>
<tr>
<th>Objectives</th>
<th>Baseline</th>
<th>Year 1</th>
<th>Year 2-6</th>
<th>Year 7/Goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Objective 1: Study the reliability of the revised IE instrument.</td>
<td>Preliminary data available for 2012 IE instrument. Unknown for new 2016 IE instrument.</td>
<td>Finalize methodology and results for 2012 IE, including percent agreement by rater types and correlation coefficient.</td>
<td>Replicate analyses with 2016 IE once sufficient sample achieved. A correlation coefficient (alpha) of .70 or higher is established by Year 5.</td>
<td>A correlation coefficient (alpha) of .70 or higher is maintained on 2016 IE.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objective 2: Improve rater training materials with an explicit focus on rater calibration</td>
<td>Content of existing training materials and methods of delivery identified.</td>
<td>Begin development of training materials for supervisors to account for 2016 IE.</td>
<td>Finalize and update as appropriate training materials for supervisors. Determine ways to assess calibration during training.</td>
<td>Target is at least 90% agreement.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Progress Since SI Plan Proposed

- Meeting of University Supervisors (December 2016) to generate list of evidence that could be collected to support rating decisions.

- Examined preliminary reliability data between university and field-based supervisors (descriptives, Kappa, Intra-class Correlations) from fall 2016 and spring 2017 administrations of IE (summer 2017).

- Initial analyses indicate fair agreement between ratings of supervisors.

- “All Programs” meeting (August 2017) to review key program information, including rating policy.
Next Steps

■ Reliability
- Continue to collect and examine reliability data
- Work to improve existing and develop new training materials for supervisors; consider delivery methods for training (e.g., online)

■ Predictive Validity
- Consider relationship between performance assessed on IE during culminating internship and measures of completer effectiveness (i.e., Standard 4), including
  - VAM
  - Teacher effectiveness ratings as part of performance evaluation