

## **5.2.2 Development and Validation Process of Instruments Used in the Quality Assurance System**

This document provides information concerning the process to construct the instruments used within the EPP quality assurance system. The data complements evidence from the 5.1.1 (EPP Quality Assurance System Model) and evidence from the 5.1.3 Data Collection and Analysis Process.

### **Section A**

#### **University Placement Testing and Admissions (PEAU)**

The University Placement Testing and Admissions (PEAU), also known as the College Board Test, are criteria of admission in accredited university institutions in Puerto Rico. They are constructed and managed by the Puerto Rico and Latin American Office of the College Board. (<https://latam.collegeboard.org/page/peau>).

#### **Test Content**

The tests consist of two main parts: the Academic Aptitude Test (Verbal Reasoning and Mathematical Reasoning) and the Academic Achievement Tests in the subjects of Spanish, Mathematics and English.

#### **Part I – Academic Aptitude Test**

**Verbal Reasoning:** measures students' verbal skills development level. It evaluates the ability to use verbal material by means of reading interpretation.

**Mathematical Reasoning:** measures students' ability to manage and apply the mathematical principles and concepts in solving problems related to arithmetic, algebra, geometry and statistics and probability.

#### **Part II – Academic Achievement Tests**

Measures specific knowledge learned on the subjects of Spanish, Mathematics and English as a second language. The scores obtained in these tests allow universities to identify students who qualify for honor or advanced courses. They also allow identifying students' academic weaknesses to be able to provide the help they need to pursue higher education.

## Section B

### Clinical Practice Evaluation Instrument

The Clinical Practice Evaluation Instrument was developed in 2003; and has been revised in several occasions. This document presents the latest validation process which concluded in 2010.

**Stage one:** During the first stage of the validation process, the competencies developed through the Conceptual Framework approved by the Faculty of the College of Education were reviewed. At that time the main changes were: Establishment of an emphasis on diversity, integrating the research component within the implementation competency, incorporating a technology competency, and the fusion of the competency related to professional performance with the professional and personal development competency.

**Stage two:** During the second stage, the revised instrument was distributed to all clinical practice professors and their feedback was included in the instrument. We received feedback from teachers at the preschool and elementary level; eight components were added to the two competencies that were fused during the first stage (these were divided into four components); the technology component was eliminated from the implementation competency - since during the first stage this component (technology) became a separate competency and in its place, the research and creation component was added; and, lastly, more manifest discrimination criteria were established between the achieved and the outstanding levels.

**Stage three:** The third stage primarily consisted of a review of the document developed during the second stage by a group of professors representing the preschool, elementary, and secondary levels; assessment; and other areas not related to primary subjects. During this phase, professors representing the different levels/areas verified the pertinence of the language used and also verified all established criteria related to their professional area. Additionally, they worked on ranking the performance levels of each component within all of the competencies.

**Stage four:** The fourth stage was the content validation of the most recent version of the instrument. During this phase we have included an alignment of the competencies included in the instrument alongside the dispositions of the future teacher as outlined within the Conceptual Framework of the EPP. Also, within this phase we included representatives from different sectors, including professors and students, cooperating teachers, and administrators.

During the validation process a template was used (See evidence 5.2.1) for the validation of the Clinical Practice Evaluation Instrument) where the 12 evaluators performed their evaluation. The template included the 10 EPP competencies and the evaluation of the following criteria:

Pertinence of the information, Content clarity, Adequate writing style, Adequate scaling of the performance levels, and their relationship with the dispositions. For each one, the evaluators used the following scale: 3: Good, 2: Average, and 1: Needs Improvement. Three was indicative that the evaluator was completely satisfied with the statement and one would indicate that this premise needed improvement because it did not comply with the established criteria. An additional column was added for comments. To complete the validation process the evaluator had the evaluation instrument, as well as the page in which the dispositions are presented.

According to the Conceptual Framework of the College of Education, dispositions are more directly related with competencies 2, 3, 5, 8, 9, and 10. It is for said reason that competencies 1, 4, 6, and 7 are marked as Not Applicable in the alignment with dispositions document.

**Results:** Evidence 5.2.2 shows that all of the competencies were evaluated by more than 70% of the experts as Good in the areas evaluated in Pertinence of the information and Adequate scaling of the performance levels. In Content clarity and Adequate writing style, only Competency 8 obtained 64% in both items. In these cases, the suggestions for corrections by the evaluators were followed.

**Reliability analysis:** In order to know if this instrument consistently measures what it is intended to measure, a reliability analysis (internal consistency) was performed using Cronbach's Alpha procedure. The analysis was performed for the 2013-2014, 2014-2015, and 2015-2016 years. Evidence 5.2.3 shows the sample sizes used in the analysis and the reliability coefficient (Cronbach's Alpha). Results from the analysis shows high reliability coefficients for these three years.

## Section C

### **Rubrics for the evaluation of the Electronic Portfolio**

The Electronic Portfolio uses rubrics to evaluate the reflections and evidences that students present to demonstrate mastery of the EPP competencies. What follows presents the process carried out by professors to calibrate the rubric to evaluate reflections and evidence presented by students.

**Process to calibrate the rubric for the Electronic Portfolio:** What follows is a description of the process followed by the professors to calibrate<sup>1</sup> the Porta-e rubric. The intention was to get the highest consensus in the interpretation [of the rubric] and to promote the validity of assessment results. 19 individuals participated in the process: Professors teaching the Reflective Seminar FAED 4001 and 4002, department directors, the director of the evaluation unit, and the assessment coordinator.

**Procedure:** Samples of evidence and reflections collected through *Mi Trayectoria* and Principles 2, 4, 7, and 10 that were being worked with at the time within the Reflective Seminar FAED 4001 and 4002. The rubric designed for said purpose was used and the statistical results among evaluators were compared.

---

<sup>1</sup> In calibration, a group of scorers all look at the same piece of work, score the work using a rubric, and then discuss how they scored the work. Disagreements within the group regarding the scores lead to focused discussions about scoring. These discussions typically include two major topics: the weight and worth of evidence from the work, and what the rubric means. As the discussion continues, problems and solutions are recorded as “calibration notes” for future scorers. Particularly clear examples of work at different levels are also kept for future scorers as benchmarks. Cited from Rhode Island Skills Commission (2005) *Protocol for Calibrating and Scoring ELA Tasks*. [http://www.ride.ri.gov/highschoolreform/dslat/pdf/por\\_100201.pdf](http://www.ride.ri.gov/highschoolreform/dslat/pdf/por_100201.pdf)

**Results obtained through the use of the rubric:** To evaluate the work submitted, the professors used the quality criteria established in the rubric to evaluate the E-Portfolio.

### Electronic Portfolio Rubric

Criteria (C)	Not evidenced 0	Initiated 1	In processs 2	Attained 3
1. The evidence selected properly illustrates the attainment of the principle/competency. (Double weight upon adding)				
2. Describes and contextualizes the evidence in a clear and precise manner.				
3. The justification explains the relationship between the evidence and the principle in a clear and convincing manner.				
4. The justification reflects the significant development of the student as it relates to their learning and realization as a teacher. (Double weight upon adding)				
5. The writing in the justification shows appropriate use of language. Ideas are presented clearly and coherently.				

The results of the calibration process were as shown below.

### Principle: #2 Learning and development

Professor	Total	C1	C2	C3	C4	C5
Sandra Macksoud López	14	2	3	2	4	3
Ana Quijano	21	6	3	3	6	3
José del Valle	21	6	3	3	6	3
Consuelo Torres Burgos	12	2	3	2	2	3

The average of the evaluation in this group was 17 out of a possible 21, which translates to 81%. The total scores granted by the evaluators ranged from 12 to 21, displaying considerable divergence among the evaluators.

### Principle: # 4 Critical thinking, research and creativity

Name	Total	C1	C2	C3	C4	C5
Annette López	17	4	2	3	6	2
Ileana Quintero	11	2	2	1	4	2
Claudia Álvarez	11	4	1	1	4	1

The average of the scores for this principle was 13 out of 21, which represents 61.90%. In the evaluation of this principle there were three evaluators and two of them granted the same score. Convergence within this principle showed that two evaluators were in agreement and provided the same score and one evaluator had a different score; this phenomenon occurred with all criteria. It could be said that convergence and divergence was two to one respectively. Divergence of opposing points were not well marked, the evaluators differed

but to a low degree, to wit, two to one and that occurred in all criteria related to principle #4.

### Principle: #7 Educational technologies

Name	Total	C1	C2	C3	C4	C5
Zulma Medina	7	2	1	1	2	1
Carmen Pacheco	21	6	3	3	6	3
Juan Meléndez	21	6	3	3	6	3

**Average:** 16.33/21 this average is close to an ideal score with 77.76%. It could improve to reach a performance of 80% or more. As situation similar to the previous principle occurred with this principle since two evaluators agreed to grant the same maximum score. The third evaluator granted very low scores equivalent to 33% of the ideal score.

### Principle: # 10 Professional and reflective professional action and development

Name	Total	C1	C2	C3	C4	C5
Anita Yudkin	17	4	3	2	6	2
Annette López	19	6	3	2	6	2
Ileana M. Quintero	17	6	2	2	6	1
Claudia Álvarez	19	6	3	2	6	2

**Average:** 18/21 this result shows that the score obtained is almost ideal with 85.7%. In this case the evaluators coincided in granting similar scores that did not differ significantly from one another. There was not much variation in the scores granted. It ranged between 17 and 19.

**Average total for all four principles** 15.89/21. This result is the sum of the averages for all four principles that were developed during the meeting to discuss the four principles. The 15.89 total average represents 75.66% of the ideal score. In this case 100% is equal to 21.

**Conclusions:** In general there was considerable convergence among evaluators, with some exceptions. In each group there was always a professor that distanced themselves from the others, and the concluding discussion of the evaluation helped to gain understanding of the way the evaluation was done by each one of the participants so as to achieve higher consensus. It is possible that differences in the backgrounds and prior experiences of some of the evaluators led to differences in the interpretation of the rubric and the evaluation of sample work. The results from this experience allowed us to see that students' performances are not reaching the desired maximum level in any of the principles.

In this sample all principles would have recommendations for the improvement of the performance evidence. A search for explanations for this phenomenon would be advised in order to determine if it is related to lack of understanding of the principles, use of technology issues, the importance students give to the development of the Porta e, the motivation students

have, the evaluation criteria professors use, since through the sample used we can see they vary as reflected by the substantial difference in scoring among evaluators.

### **Recommendations:**

1. To hold a second, follow-up meeting to present results and exchange impressions and actions needed to review and improve the Seminar, the Porta e, the rubrics and checklists, and the evaluation process.
2. Wide dissemination of the results of this meeting so that future Porta e evaluators can take into consideration the aspects identified to promote consistency and equity in the evaluation of the Porta e, and thus they can reflect on their practice and can provide the means to improve these processes.
3. Hold a meeting to discuss and analyze the principles and reflections that were not included in this initial process in order to complete the consistency test among all evaluators using the Porta e and everything else related to these principles and reflections.

## **Section D**

### **Exit Survey**

#### **Objectives**

The purpose of the Exit Survey Questionnaire is to collect candidates' data at the third transition point in order to know the competencies and dispositions they acquired through the EPP preparation. We use the data collected to do correlations with the Clinical Practice Evaluation Instrument data to determine candidates' perception about the preparation program.

#### **What is measured**

It measures candidates' perceptions about the competencies and dispositions they acquired through the EPP preparation.

#### **Alignment**

The instrument is aligned with EPP competencies and dispositions, CAEP standards, and InTASC standards.

#### **Instrument construction process**

The construction of the instrument was carried out by the Educational Research Center of the EPP. Using a four level Likert scale, students were asked about their knowledge or mastery of the professional competencies and the dispositions included within the College's conceptual framework.

#### **Administration**

The instrument is administered every semester during the last day of the Clinical Experience Course.

### **Analysis**

Data is processed by personnel attached to the Educational Research Center using SPSS.

## **Section E**

### **Technology Survey**

#### **Objectives**

- Identify the integration of technologies by professors is promoted among the candidates.
- Identify activities professors promote among candidates to present within the electronic portfolio.
- Use data generated to align technology activities promoted by the professor with the competencies so these also align with CAEP standards.

#### **Alignment**

The instrument is aligned with technology aspects within the EPP competencies, CAEP standards, and InTASC standards.

#### **What is measured**

- The different technologies used by professors in their courses which are geared to the promotion of their integration/adoption among the candidates.
- Technology integration/adoption activities within the classroom.

#### **Instrument construction process**

- The ARTI Department was in charge of its construction.
- A preliminary questionnaire was designed based on the stated objectives and an evaluation of that instrument was carried out by technology professors.
- The instrument was then revised by the Academic Affairs Office.
- Later on it was reviewed by program directors, the director of Educational Research Office (CIE for its acronym in Spanish), and the director of the Evaluation Office.

#### **Administration**

It was administered for the first in October 2016. It was answered by 57% of the professors. It was administered online.

**Analysis**

The ARTI Department was in charge of the analysis and interpretation.

**Section F****Completers Satisfaction Survey****Objective**

The main objective is to determine completer's satisfaction with the preparation they received through the EPP.

**What is measured?**

The satisfaction questionnaire is divided into four sections: A) Academic and employment data, B) Competencies, Dispositions, and Professional Development, C) Satisfaction with UPR's teacher preparation program, and D) Evidence of their Educational Practice. Section A has 17 questions in which graduates describe their current employment situation, their academic background, and the local environment. Section B has 10 competencies, with three or four items for each one, the graduate must pick the level of competence they have. The scale has four possible answers ranging from highly competent to not competent. Section two also include five questions concerning dispositions and 14 questions about areas of professional development. Section C includes 12 items related to the level of satisfaction graduates have for UPR's teacher preparation program. Section D has four questions concerning the graduate's practice in the school and in their work environment.

**Alignment**

The questionnaire is aligned with the requirements established by CAEP for Standard 4 and the InTASC standards.

**Instrument construction process**

The instrument was constructed using as guides the instruments developed by the Academic Planning Office of UPRRP, the Teaching Practice Office, the Evaluation Office, and the Center for Educational Research. To ascertain the content validity of the instrument a team of six experts reviewed the instrument. A pilot test was carried out as well.

**Administration**

It was administered digitally using Google Forms.

**Analysis**

The director of the Induction Office was in charge of analyzing the data along with her research assistants. A reliability analysis (internal consistency) was performed using Cronbach's Alpha. Results from this analysis showed high (.98) reliability coefficients.

## Section G

### School Director's Opinion Survey

#### **Objectives**

Learn the opinion of school directors concerning the performance of EPP completers as related to professional achievements and student achievement.

#### **What is measured**

The questionnaire for school directors has 29 questions out of which eight are sociodemographic questions geared toward collecting information about the location and type of jobs held by graduates. The 12 main questions are related to the competencies and dispositions they display and master as graduates of the College of Education, UPR-RP. There is one question concerning the educational philosophy displayed by the teachers and eight about the directors' opinion of the graduates with an emphasis on the teachers' professional achievements, the achievements reached by their students, and their opinion concerning the quality of the teacher preparation program in general.

#### **Alignment**

The questionnaire is aligned with the requirements established by CAEP for Standard 4 and the InTASC standards.

#### **Instrument construction process**

The instrument was constructed using as guides the instruments developed by the Academic Planning Office of UPRRP, the Teaching Practice Office, the Evaluation Office, and the Center for Educational Research. A review was carried out of the instrument developed in 2007 and revised in 2013 and 2015 by the Center for Educational Research. To ascertain the content validity of the instrument a team of six experts reviewed the instrument. A pilot project that included 28 school directors was also carried out.

#### **Administration**

It was administered through a digital link that used Google forms.

#### **Analysis**

The director of the Induction Office was in charge of analyzing the data along with her research assistants.