EDUCATION POLICY INITIATIVE *at* CAROLINA

Initial Findings from edTPA Implementation in North Carolina

Since 2010, select teacher preparation programs (TPPs) in the UNC system have used edTPA, a teacher candidate performance assessment developed by Stanford University, to help determine teacher candidates' readiness to teach and to provide valid and reliable assessment data for program evaluation and continuous program improvement. In this policy brief we summarize the history of edTPA implementation in North Carolina, show how UNC system candidates are scoring on edTPA, examine whether edTPA scores measure the instrument's three main teaching tasks, and assess whether edTPA scores predict the performance of first-year teachers. Key results indicate (1) benefits to the phased and gradual implementation of edTPA in North Carolina; (2) that a majority of teacher candidates have an acceptable level of knowledge and skills to begin teaching; and (3) that edTPA scores predict the value-added estimates and evaluation ratings of first-year teachers. These results have implications for continued edTPA implementation in North Carolina, setting edTPA cut scores for high stakes teacher licensure decisions, and using edTPA data for TPP accountability and improvement.

Introduction

In recent years policymakers and accreditation agencies have strengthened regulations holding teacher preparation programs (TPPs) accountable for the performance of their graduates and have encouraged TPPs to use evidence for continuous program improvement. For many teacher educators, one response to this data-driven context has been support for the creation and widespread adoption of teacher candidate performance assessments. Candidate performance assessments are portfolios completed by teacher candidates during their student teaching experience that consist of curriculum plans, video clips of instruction, student work samples, and candidates' reflective commentaries. States and TPPs can use these portfolios to determine candidates' readiness to teach—linking candidate scores to teacher licensure decisions. Perhaps more importantly, given the ability of performance assessments to identify program strengths and areas of concern, TPPs can use performance assessments as a basis for evidence-based improvement.

In North Carolina, the UNC system and its constituent institutions are leading the push to integrate candidate performance assessments into teacher preparation. Specifically, UNC system institutions have adopted edTPA, a candidate performance assessment developed by Stanford University and used by over 700 TPPs in 41 states. In this policy brief, we detail the history of edTPA implementation in North Carolina, show how UNC system teacher candidates are scoring on edTPA, and assess the validity of edTPA scores—do they measure the instrument's key teaching constructs and predict graduates' performance as beginning teachers. For TPPs and policymakers, both in North Carolina and nationwide, we hope this evidence benefits considerations of edTPA adoption, setting cut scores for teacher licensure decisions, and using edTPA scores for program improvement.

Background

The edTPA is comprised of three main teaching tasks—Planning, Instruction, and Assessment—with five scored rubrics within each task.[1] Evaluators score each rubric from 1 to 5, with a 1 indicating a struggling candidate who is not ready to teach, a 2 indicating a candidate who needs more practice, a 3 indicating an acceptable level of performance to begin teaching, a 4 indicating a candidate with a solid foundation of knowledge and skills, and a 5 indicating a highly accomplished teacher candidate. Teacher preparation programs can score these portfolios locally, with trained faculty and university supervisors, or submit the portfolios for national scoring by officially-calibrated evaluators. To cover teacher candidates across a range of licensure areas, there are 26 different edTPA content areas (e.g., elementary literacy, middle childhood science, secondary history/social studies, special education, world languages, etc.).

In this policy brief, all edTPA scores come from national scoring. To assess how UNC system teacher education candidates are scoring on edTPA, we use data from East Carolina University (ECU) for 2013-14 through 2015-16, UNC Charlotte (UNCC) and North Carolina State University (NCSU) for 2014-15 and 2015-16, and UNC Chapel Hill (UNCCH) for 2015-16. We use these same data to examine the construct validity of the national scores. To assess the predictive validity of edTPA scores, we use data from ECU's 2013-14 graduating cohort and teacher performance data (EVAAS estimates and NCEES ratings) from the 2014-15 school year. With teacher performance data from 2015-16, we will soon be able to assess the predictive validity of edTPA scores for graduates of multiple institutions.

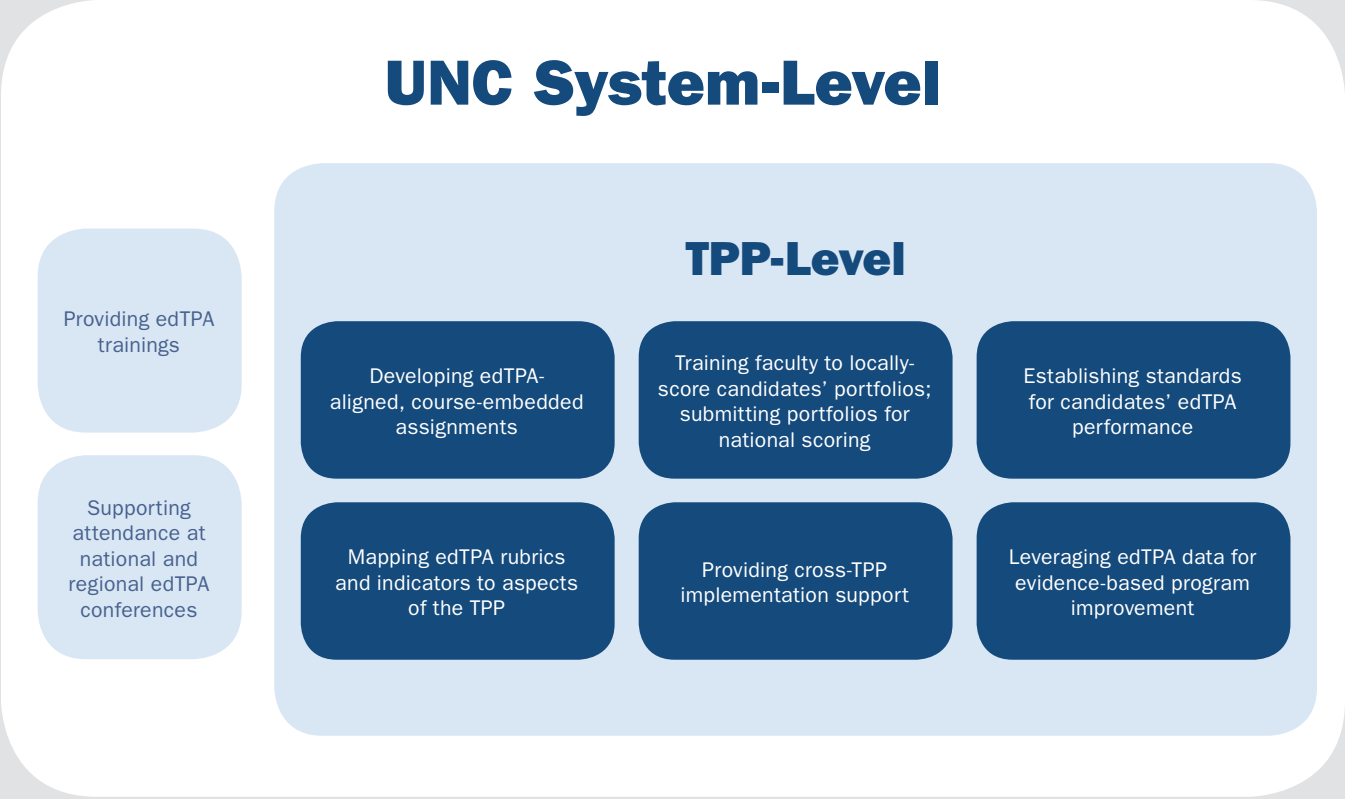
What progress have teacher preparation programs made in implementing edTPA?

Teacher preparation programs in North Carolina first engaged with edTPA, then the Teacher Performance Assessment Consortium (TPAC), in 2010. Initially, education deans at ECU, UNCCH and Winston-Salem State University led efforts to bring the performance assessment to their TPPs. Later, edTPA was expanded through a pilot supported by the UNC General Administration to include these original TPPs and NCSU, UNCC, UNC Asheville, and Western Carolina University. In 2015, following the UNC Board of Governors recommendations for improving teacher and school leader preparation, all TPPs in the UNC system were invited to engage in a system-wide edTPA implementation. The UNC system is now working with each member institution to (1) assess readiness for edTPA implementation or successes/challenges with implementation thus far and (2) develop plans for edTPA implementation and growth. Additionally, several private and independent TPPs in North Carolina are exploring edTPA as a program option.

Overall, edTPA implementation in North Carolina has progressed through action at the UNC system and individual TPP levels. As shown in Figure 1, actions at the UNC system level include financial supports for edTPA training and attendance at national and regional edTPA conferences. Actions at the TPP level include developing course-embedded assignments aligned with edTPA, training faculty and university supervisors to locally-score edTPA portfolios, submitting candidates' portfolios for national scoring, establishing standards for candidates' edTPA performance (for high-stakes decision-making at the TPP level), mapping edTPA rubrics and indicators to aspects of the TPP, providing cross-TPP implementation support, and leveraging edTPA data for evidence-based program improvement. Reflecting their diverse missions, cultures, and contexts, UNC system TPPs have engaged in these activities in different ways and in different combinations. For example, ECU has been a strong proponent of local-scoring, setting internal cut score requirements for awarding licensure, and using edTPA data for program improvement. Other programs, such as NCSU and UNCC, have relied on national edTPA scoring and have not fully instituted cut score requirements for candidate licensure. Importantly, the phased and gradual edTPA implementation in North Carolina has provided TPPs time and space to support each other and to develop a growing edTPA teacher educator community.

[1] The edTPA Elementary Education handbook includes 18 rubrics; the edTPA World and Classical Language handbooks include 13 rubrics.

Figure 1: edTPA Implementation Activities and Supports in North Carolina



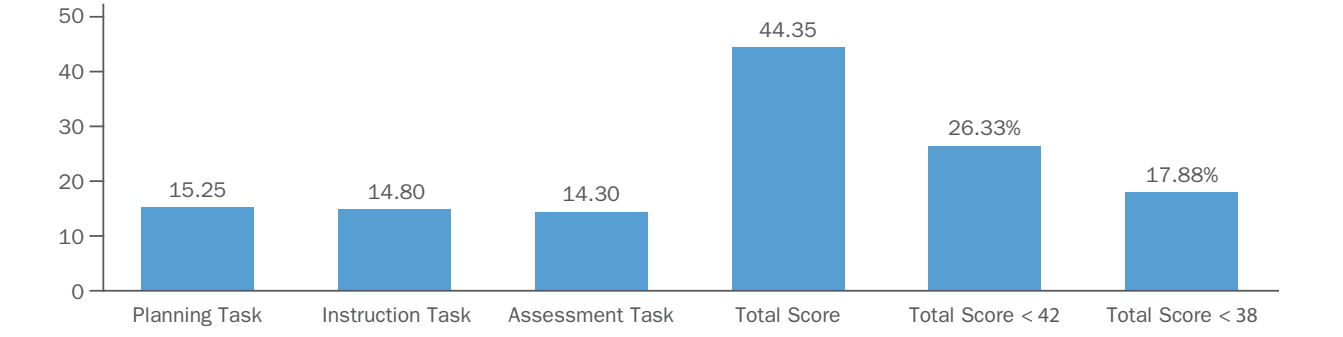
How are teacher candidates scoring on edTPA?

To assess how UNC system teacher education candidates are scoring on edTPA, we use all available national scores for ECU, NCSU, UNCC, and UNCCH graduates. Here, Figure 2 illustrates that, on average, UNC system candidates have an acceptable level of knowledge and skills to begin teaching. Specifically, the average edTPA total score is 44.35, while the average scores on the three edTPA tasks of Planning, Instruction, and Assessment are 15.25, 14.80, and 14.30,[2] respectively. The right two bars in Figure 2 show that approximately 26 percent of candidates had a total score below a proposed edTPA cut-score of 42; nearly 18 percent of candidates had a total score below a proposed cut-score of 38. While this evidence is important as North Carolina considers an official cut-score (or cut-score range) for teacher licensure, we note that most of the candidates in our sample did not have to meet a minimum total score to be recommended for licensure by their respective institutions. In a high stakes setting, where candidates need a certain score for licensure, it is possible that fewer candidates will score below these thresholds.

Evidence of teacher candidates' readiness to enter teaching is better shown in Figure 3, which displays the percentage of teacher candidates scoring at each level of the 15 edTPA rubrics. The majority of teacher candidates score at a level three or higher on all 15 edTPA rubrics—level three represents the knowledge and skills of a candidate who is ready to teach. Individual rubric scores reflect TPP strengths in *Planning for Content Understanding*, the classroom *Learning Environment*, and *Providing Feedback to Guide Further Learning*. The lowest scoring rubric is *Student Use of Feedback*, where only 60 percent of teacher candidates scored at a level three or higher. Overall, these rubric scores mirror scores from TPPs across the United States, indicating that teacher candidates may require additional practice in the area of assessment, particularly with opportunities to guide *Students' Use of Feedback*.

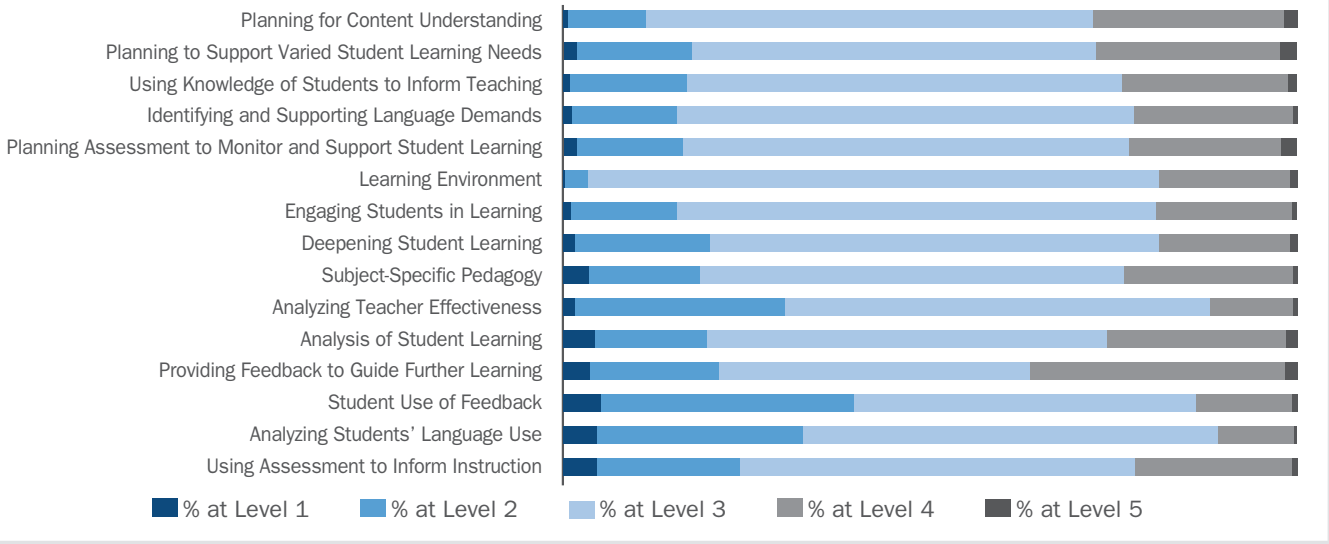
[2] A standard-setting process led by three panels of educators and policy makers resulted in a recommended cut-score band ranging from a total score of 37-42. http://edtpa.aacte.org/faq#51

Figure 2: Summative Values for edTPA Scores



Note: This figure displays the average *Planning*, *Instruction*, *Assessment* and total score as well as the percentage of candidates with a total score below 42 and below 38.

Figure 3: The Distribution of edTPA Scores



Note: This figure displays the distribution of edTPA rubric scores. For these distribution values, we rounded rubric scores with a decimal value (e.g. 2.5 or 3.5) up to the nearest whole number.

Does edTPA measure the instrument's main teaching tasks?

For construct validity analyses we use all of the national scores currently available in North Carolina—from ECU for 2013-14 through 2015-16, from UNCC and NCSU for 2014-15 and 2015-16, and from UNCCH for 2015-16—to assess whether portfolio scoring identifies edTPA's three main teaching tasks. Specifically, we use parallel analysis to determine the number of latent constructs in the data; our exploratory factor analysis approach allows the latent constructs to be correlated with each other since they measure components of an integrated teaching process. Results in Table 1 show that our analyses revealed a three factor structure for the nationally-scored edTPA portfolios. All five *Planning* rubrics load onto a single factor, four of five *Instruction* rubrics load onto a single factor, and all of the *Assessment* rubrics (plus one *Instruction* rubric) load onto a single factor. These results replicate the factor structure found in the edTPA field test report—there *Analyzing Teacher Effectiveness* also loaded with the *Assessment* task rather than the *Instruction* task.[3]

[3] Factor analysis of locally-scored edTPA portfolios also reveals a three factor structure with the *Analyzing Teacher Effectiveness* rubric loading with the *Assessment* task.

Table 1: Does edTPA measure the instrument's main teaching tasks?

Task	edTPA Rubric	Factor 1	Factor 2	Factor 3
Planning	Planning for Content Understanding	**0.788**	0.124	-0.131
	Planning to Support Varied Student Learning Needs	**0.781**	-0.017	0.018
	Using Knowledge of Students to Inform Teaching	**0.663**	-0.008	0.163
	Identifying and Supporting Language Demands	**0.718**	-0.030	0.101
	Planning Assessment to Monitor and Support Student Learning	**0.787**	0.000	0.020
Instruction	Learning Environment	-0.003	**0.838**	-0.124
	Engaging Students in Learning	0.101	**0.740**	0.056
	Deepening Student Learning	0.001	**0.779**	0.086
	Subject-Specific Pedagogy	-0.051	**0.720**	0.140
	Analyzing Teacher Effectiveness	0.217	0.120	**0.436**
Assessment	Analysis of Student Learning	0.135	0.061	**0.663**
	Providing Feedback to Guide Further Learning	-0.091	0.042	**0.803**
	Student Use of Feedback	-0.091	0.002	**0.861**
	Analyzing Students' Language Use	0.150	0.031	**0.611**
	Using Assessment to Inform Instruction	0.161	-0.005	**0.684**
	Cases	2153		

Note: This table presents factor loadings for the nationally-scored edTPA portfolios from ECU (2013-14 to 2015-16), UNCC (2014-15 and 2015-16), NCSU (2014-15 and 2015-16) and UNCCH (2015-16). All factor loadings greater than 0.40 are bolded.

Do edTPA scores predict beginning teacher performance?

For our predictive validity analyses, we focus on members of ECU's 2013-14 graduating cohort who became first-year teachers in North Carolina public schools in the 2014-15 school year. To assess relationships between edTPA measures and teacher value-added, we use teachers' standardized EVAAS estimates (standardized across all teachers in North Carolina and within End-of-Grade/End-of-Course/final exam) and run linear regression models. To examine relationships between edTPA measures and teacher evaluations, we use teachers' ratings on the five NCEES standards directly assessed by school principals (Leadership, Classroom Environment, Content Knowledge, Facilitating Student Learning, and Reflecting on Practice) and estimate ordered logistic regression models.[4] Each of these analyses controls for a limited set of school characteristics. For these teacher performance measures we estimate three sets of models: (1) *Task* models for the edTPA Planning, Instruction, and Assessment constructs (identified by factor analysis); (2) *Rubric* models for each of the 15 edTPA rubrics (entered individually into models); and (3) *Summative* models for the edTPA total score (standardized) and an indicator if the candidate had a total score of 42 or above. As one limitation, we note that these predictive validity results are for one graduating cohort from one institution—results with additional cohorts and institutions may differ.

[4] Principals can rate teachers at one of five performance levels—not demonstrated, developing, proficient, accomplished, and distinguished. For more information on the NCEES, please see the following: http://www.ncpublicschools.org/docs/effectiveness-model/ncees/instruments/teach-eval-manual.pdf

Table 2: Do edTPA Scores Predict Teacher Performance?

edTPA Measures	Teacher Value-Added	Teacher Evaluation Ratings (NCEES)				
	Std. EVAAS	Leadership	Classroom Environment	Content Knowledge	Facilitating Student Learning	Reflecting on Practice
Planning Factor	0.080	0.862	**0.619***	0.794	0.723	0.746
Instruction Factor	0.086	**1.664****	**1.828****	1.407	**1.941****	**1.621***
Assessment Factor	0.090	**1.611****	1.424	1.248	**1.494+**	1.201
Planning for Content Understanding	**0.237***	**1.686***	1.436	1.283	1.341	1.226
Planning to Support Varied Student Learning Needs	**0.183+**	**1.732***	1.166	1.133	1.374	1.565
Using Knowledge of Students to Inform Teaching	0.184	0.956	0.916	0.941	1.168	0.640
Identifying and Supporting Language Demands	0.130	**2.395****	1.197	1.290	**1.831***	1.032
Planning Assessment to Monitor and Support Student Learning	**0.218+**	1.194	0.758	0.783	0.963	0.958
Learning Environment	0.096	**2.151***	**2.065+**	1.700	**2.815****	1.502
Engaging Students in Learning	0.106	**3.829****	**2.466***	**1.996+**	**2.626****	**1.987+**
Deepening Student Learning	**0.255+**	**2.291****	**2.259****	1.327	**2.679****	**1.814+**
Subject-Specific Pedagogy	**0.340***	**1.629***	1.538	1.493	**2.024****	**1.596***
Analyzing Teacher Effectiveness	**0.274+**	**2.971****	1.568	**2.017+**	**2.694****	1.438
Analysis of Student Learning	0.131	1.348	1.122	0.974	1.410	0.954
Providing Feedback to Guide Further Learning	0.154	1.401	1.185	0.912	1.318	1.206
Student Use of Feedback	**0.181+**	**2.121****	**1.725***	1.514	**1.695***	1.435
Analyzing Students' Language Use	0.196	**2.011****	**1.802***	1.441	**1.644***	1.449
Using Assessment to Inform Instruction	0.034	**1.723***	1.392	1.433	**1.618***	1.024
Standardized Total Score	**0.205***	**1.899****	**1.401+**	1.278	**1.718****	1.288
Scoring at 42 or Above	**0.593***	**3.306***	**3.515***	2.607	**3.431****	1.263
Cases	202	169	169	169	169	169

Note: This table displays regression coefficients (EVAAS) and odds ratios (NCEES ratings) for the relationship between edTPA measures and first-year teacher performance. Odds ratios greater than '1' are positive; those less than '1' are negative. +, *, and ** indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively.

Overall, Table 2 shows that edTPA scores significantly predict first-year teacher performance. Concerning teacher value-added, the standardized EVAAS column indicates that seven edTPA rubrics predict higher value-added estimates. For example, a one point increase on the *Subject-Specific Pedagogy* rubric is associated with a 34 percent of a standard deviation increase in teacher effectiveness. Summatively, the standardized total score and having a total score of 42 or greater also predict significantly higher value-added estimates.

Concerning teacher evaluation ratings, the *Instruction* factor predicts significantly higher ratings on four standards while the *Assessment* factor predicts higher ratings on two standards. To ease interpretation of these results, we converted the odds ratios to predicted probabilities. For the Facilitating Student Learning standard, beginning teachers whose *Instruction* factor score is one standard deviation below the mean have a 9.48 percent predicted probability of rating at developing and a 6.75 percent predicted probability of rating at accomplished; by comparison teachers whose *Instruction* factor score is one standard deviation above the mean have a 2.70 and 21.25 percent predicted probability of rating at developing and accomplished, respectively. Many edTPA rubrics, particularly in the *Instruction* task, predict significantly higher evaluation ratings. Conceptually, the significant results for the *Instruction* rubrics on the Facilitating Student Learning standard emphasize the alignment between teacher actions and competencies that comprise this evaluation standard and *Instruction* task indicators. Lastly, the two summative measures—the standardized total score and scoring at 42 or greater—predict significantly higher evaluation ratings for three teaching standards.

Discussion

Rather than mandate edTPA implementation, the UNC system has taken a measured and supporting role, allowing UNC system institutions to engage with edTPA of their own initiative and encouraging TPPs to collaborate with each other as edTPA spreads throughout the system. Now, edTPA is building momentum in North Carolina, with the UNC Board of Governors recommending the use of valid and reliable candidate performance assessments, some private universities considering its adoption, and the General Assembly and State Board of Education discussing candidate performance assessments as a teacher licensure requirement. Into this context, our policy brief makes several contributions. First, we show that after a few years of implementation, edTPA is becoming fully-immersed into the early-adopting TPPs in North Carolina. Knowing that it takes time for edTPA to become part of the TPP-culture is important to the potential use of edTPA in high-stakes teacher licensure decisions and perhaps, more importantly, to the use of edTPA for evidence-based TPP improvement efforts. Second, portfolio scores indicate that a majority of candidates have an acceptable level of knowledge and skills for beginning teaching. If North Carolina wants to use edTPA scores for teacher licensure decisions, we suggest that the state implement a graduated cut-score.[5] This graduated process will give TPPs without sufficient exposure to edTPA time to fully integrate the instrument into course and fieldwork and better help their candidates succeed. Beyond high-stakes licensure decisions, we also suggest that North Carolina find ways to emphasize the use of edTPA for program improvement. Third, we confirm that national edTPA scores are measuring the instrument's three main constructs of teaching practice—Planning, Instruction, and Assessment. Lastly, we provide nascent evidence that national edTPA scores predict the value-added estimates and evaluation ratings of first-year teachers. More data is needed here—from additional universities and graduating cohorts—but these results suggest that edTPA is a valid outcome upon which to base candidate licensure and program improvement decisions.

For more research on this topic

Bastian, K.C., Henry, G.T., Yi, P., & Lys, D. (2016). Teacher candidate performance assessments: Local scoring and implications for teacher preparation program improvement. *Teaching and Teacher Education*, 59, 1-12.

Goldhaber, D., Cowan, J., & Theobald, R. (2016). Evaluating prospective teachers: Testing the predictive validity of the edTPA. *Calder Working Paper 157*. Available from: http://www.caldercenter.org/sites/default/files/WP%20157.pdf

Peck, C.A., Singer-Gabella, M., Sloan, T., & Lin, S. (2014). Driving blind: Why we need standardized performance assessment in teacher education. *Journal of Curriculum and Instruction*, 8(1), 8-30.

Sato, M. (2014). What is the underlying conception of teaching of the edTPA? *Journal of Teacher Education*, 65(5), 421-434.

SCALE. (2013). *2013 edTPA field test: Summary report*. Available from: https://secure.aacte.org/apps/rl/res_get.php?fid=827&ref

[5] For example, on 15 rubric edTPA portfolios, Illinois set an initial cut-score of 35 that is rising to 37, 39, and 41 over successive graduating cohorts.

Study Authors: Kevin C. Bastian and Diana Lys (September 2016)

EPIC is an interdisciplinary team that conducts rigorous research and evaluation to inform education policy and practice. We produce evidence to guide data-driven decision-making using qualitative and quantitative methodologies tailored to the target audience. By serving multiple stakeholders, including policy-makers, administrators in districts and institutions of higher education, and program implementers we strengthen the growing body of research on what works and in which context. Our work is ultimately driven by a vision of high quality and equitable education experiences for all students, and particularly students in North Carolina.

http://publicpolicy.unc.edu/epic-home/

THE UNIVERSITY *of* NORTH CAROLINA *at* CHAPEL HILL

UNC COLLEGE OF ARTS & SCIENCES