Bridging Pre-Service Training and Professional Practice: An Analysis of the Performances of NCE Graduates in ENG 212 and In-Service Competencies

Blessing Temitope DICKSON-OMOGOYE (Ph.D)

Department of Counselling Psychology, Bamidele Olumilua University of Education, Science and Technology, Ikere-Ekiti

[ORCID ID: 0000-0002-3455-1420]

DOI: https://doi.org/10.51244/IJRSI.2026.1305000095

Received: 28 April 2026; Accepted: 04 May 2026; Published: 01 June 2026

ABSTRACT

This study examines the predictive relationship between pre-service academic performance and in-service teaching competence among Nigeria Certificate in Education (NCE) graduates, focusing on ENG 212 (Applied English Linguistics). Anchored in systems theory and assessment validity theory, the research evaluates whether pre-service assessments effectively identify future teaching competence. A concurrent explanatory mixed-method design was employed, involving 60 graduates from Federal Colleges of Education in southwestern Nigeria. Quantitative data were derived from academic records and structured classroom observations, while qualitative insights were obtained through interviews and field notes.

Findings revealed a weak and statistically insignificant correlation between pre-service performance and in-service competence ($r = 0.214$, $p = 0.132$), indicating limited predictive validity. Chi-square analysis also showed no significant association between categorical performance levels. Qualitative evidence highlighted notable divergence, with some low-performing graduates demonstrating strong teaching competence and some high-performing graduates performing poorly in practice.

The study concludes that current assessment practices, which emphasize theoretical knowledge, fail to capture essential teaching competencies. It recommends a shift toward performance-based assessment, strengthened practicum experiences, and longitudinal validation studies to enhance the effectiveness and credibility of teacher education assessment systems.

Keywords: Pre-service training, In-service competence, Assessment validity, Predictive validity, NCE graduates, Teacher education, Nigeria.

INTRODUCTION

Background to the Study

The world-wide systems of teacher certification have a fundamental assumption, which is that pre-service assessment can reveal those who will be effective professional teachers, telling the difference between those who have acquired the necessary competencies and those who have failed to do so and consequently including only those who are sufficiently prepared to receive certification to join the teaching profession (Darling-Hammond, et.al. 2020). This assumption underlies the entire rationale of teacher education as a gatekeeping profession: students must complete specified programmes of study and evaluation, and those who demonstrate satisfactory performance are awarded certification qualifications; this certification is intended to signal to employers and the general population that certified individuals possess the knowledge, skills, and dispositions necessary for competent practice. However, we must scientifically study this fundamental belief rather than accept it on faith. As long as the performance in the pre-service assessment is not predictive of future

professional teaching competence, then the whole certification system is deprived of the predictive validity that it needs to serve its intended social purpose of assuring teaching quality (Zeichner & Conklin, 2008).

The Nigeria Certificate in Education (NCE) is the lowest professional teaching qualification needed in order to be employed as a teacher in a primary and junior secondary school as required by the National Policy on Education, since it was revised in 2004. The NCE programme is a three-year postsecondary degree programme that includes general education courses, subject-matter specialisation courses, and courses in professional education that include approaches to pedagogy and learning theory and practicum experiences where pre-service teachers are supervised to apply their emerging knowledge and skills to real classroom situations. During this preparation time, the students will be assessed and periodically tested through a series of assessments which are cumulative to the student and which will help ascertain whether or not he or she has attained programme standards and thus qualifies to be certified. These examinations usually involve written tests that examine theoretical knowledge; tasks that involve the application of concepts; practical tests of teaching abilities in practicum; and comprehensive final exams. These diverse examinations successfully result in the NCE certification, which in theory assures that the candidate has been able to demonstrate the competencies of successful teaching (NCCE, 2012).

Yet, it is a certification process that supposes that the tests given in the pre-service training effectively measure the competencies that separate the effective and ineffective teachers and that the performance in the tests is a predictor of the future competence in the profession. The theory of assessment validity, as developed by Samuel Messick in his classic 1989 chapter on validity in the measurement of education, determines that validity is not a characteristic of assessment measures but the appropriateness, meaningfulness, and usefulness of the particular inferences made when using assessment scores to serve specific purposes. In the case of teacher education, the most important inference is that the performance of pre-service teachers reflects their readiness to be working in professional practice and forecasts their future effectiveness as teachers. (Grossman, et.al. 2009) This is a predictive validity claim, which can and must be tested empirically, by analysing the degree to which pre-service performance is related to measures of professional teaching competence.

Although this validity question is fundamentally significant to the overall field of teacher education and to NCE certification, in particular, strikingly little research has investigated the empirical linkage between pre-service training performance and eventual professional teaching competence in the Nigerian setting. Although many studies have explored the different features of teacher preparation, such as analyses of what is in the curriculum, how practicum experiences are, surveys of graduate satisfaction, and how well their employers view them, very few studies have directly tested the predictive validity assumption by following specific cohorts of graduates from pre-service to systematic measurement of their professional competence in classroom practice. This gap in research leaves unanswered the basic question of whether the current assessment practices effectively identify teaching capability or only assess academic performance on tasks that may or may not be significantly related to professional effectiveness (Adeleke and Dickson-Omogoye, 2023).

The Validity Question in Teacher Education

The concept of validity in educational and psychological measurement has evolved significantly from early views that regarded it as a mere property of tests determined by various forms of validity evidence (content validity, criterion validity, and construct validity) to present-day unified perspectives that interpret validity as primarily concerning the correctness of inferences and actions based on assessment scores. The most powerful reformulation by Messick (1989) conceptualised validity as a combined evaluative statement concerning the extent to which empirical data and theoretical justifications confirm the sufficiency and suitability of inferences and actions that are based on test scores or any other form of evaluation. The given definition emphasises that validity is not a fixed concept, but rather a process of gathering evidence to justify specific interpretations and uses of assessment results (Singh, 2014).

In high-stakes certification tests in teacher education, the main conclusion is that performance above specified cut-scores indicates readiness for professional practice and predicts adequate or better professional performance. This predictive claim of validity needs to be empirically supported by showing that individuals who do well on pre-service assessments, later on, demonstrate higher teaching competence than individuals

who do not do well on the same assessments. In the absence of this evidence, pre-service assessment-based certification decisions are not empirically grounded and may end up certifying ineffective people and failing to certify otherwise successful people. Social implications of invalid certification are also immense: students in the schools with underprepared teachers receive lesser-quality education, parents and communities cannot be sure that certified teachers really have the competencies required, and the profession of teacher loses its credibility and status when certification does not guarantee competence (Dickson-Omoyoye and Kehinde-Dada 2026).

The theory of assessment validity singles out several streams of evidence that can be used to determine validity. Content validity evidence pertains to the appropriate sampling of the domain of knowledge and skills that the assessment aims to evaluate. In the case of teacher education, this process will involve showing that assessments will sample the entire spectrum of teaching competencies, not focusing on certain dimensions at the expense of others (Campos, et.al. 2017). Criterion validity evidence looks at correlations between the assessment scores and external criteria of the construct of interest. The relevant type of criterion validity that applies to teacher certification is predictive validity, which analyses the extent to which the scores of an assessment are able to predict future performance on relevant criterion measures. In the case of teacher education, professional teaching competence is the end-result criterion, and hence, the correlation between pre-service assessment and future teaching performance is the most important validity evidence. The evidence of construct validity determines whether the assessments accurately measure the theoretical constructs they claim to assess and whether the interpretation of the scores aligns with established theoretical knowledge of those constructs (Roloff, et.al. (2020).

Following Messick, Kane (2006) has described an argument-based approach to validation that views validity as the use of explicit validity arguments, which define the inferences and assumptions underlying the interpretation of scores, as well as the types of evidence necessary to support or refute those assumptions; the overall argument should be accumulated to assess its plausibility. To certify teachers based on their performance in NCE programme tests, the validity argument should specify the following points: The performance on NCE programme tests measures teaching-relevant knowledge and skills (measurement assumption); those who demonstrate high levels of knowledge and skills in preparation will demonstrate high levels of teaching competence in practice (generalisation assumption); certification decisions should be based on pre-service performance (extrapolation assumption). All the assumptions must be empirically supported, and predictive validity evidence directly tests the critical generalisation assumption.

Statement of the Problem

The fall in the number of students registered in NCE English language education programmes recently, at the same time, as an ever-growing chorus of complaints by school administrators and education authorities about the poor professional performance of many NCE graduates on joining the teaching service, suggests immediate concerns about the effectiveness of the teacher preparation programmes in developing teaching competence and the reliability of the systems used to certify them as competent teachers. When pre-service tests are not effective at predicting who will become an effective teacher, this indicates a fundamental failure of validity with dire implications for the quality of education.

In particular, although colleges of education administer various tests during the NCE programme and base their certification decisions on performance in those tests to a large degree, little empirical data is available on the issue of whether pre-service performance is in fact a predictor of future professional teaching competence. The lack of evidence supporting predictive validity is particularly concerning because assessments in most colleges may rely heavily on theoretical knowledge, which is evaluated through written examinations, rather than on practical teaching competencies that should be assessed through performance in real-life situations. When tests are mostly tests of theoretical knowledge of linguistics, pedagogy or child development but do not adequately test whether the knowledge is effectively applied in real teaching contexts, high scores can reflect academic success without always reflecting teaching success.

This research addresses the critical validity gap by systematically assessing the professional teaching competence of NCE graduates in real classroom practice, based on their reported performance in ENG 212

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue V May 2026

(Applied English Linguistics), which is one of the core courses in English Language Education programmes. This research evaluates the basic validity assumption on which NCE certification is based, which states, pre-service performance is a predictor of professional competence, by empirically measuring the relationship between pre-service performance and in-service teaching competence.

LITERATURE REVIEW AND THEORETICAL FRAMEWORK

Theoretical Framework: Systems Theory and Validity Theory

This research combines two mutually supportive theoretical frameworks that, when combined, explain various facets of the linkage between pre-service training and professional practice. Systems theory, as described by von Bertalanffy (1956) and implemented in education by a variety of authors, views education as a complex adaptive system with interconnected parts (such as inputs (student selection and characteristics), processes (instruction, learning, and assessment), and outcomes (graduate competencies and credentials). One of the essential concepts of systems thinking is that successful systems are characterised by coherence and harmony among their parts: inputs must be relevant to desired outputs, processes must turn inputs into desired outputs and feedback mechanisms must allow the system to correct itself when outputs are below expectations.

Applied to teacher education, systems theory holds that when a system is operating well, there must be observable relationships between important elements within it. Professional competence (the final product) should be predicted by pre-service assessment (a process component), as they both should indicate the same underlying construct of teaching capability. When there is no relationship or a weak relationship between pre-service performance and professional competence, then this situation is a sign of system malfunction: pre-service assessment is not assessing the right things, professional practice requires competencies that have not been acquired during preparation, or both. In this sense, no or poor relationship between pre-service and in-service performance is a failed feedback loop – the system produces performance measures by way of assessment, but these measures do not reliably measure the criterion of interest (teaching competence), so the system fails in its quality assurance role (OECD, 2019).

The second theoretical framework is Assessment Validity Theory, specifically the unified conception of Messick (1989). Messick argued that validity is a unified concept encompassing all evidence bearing on whether assessment-based inferences are appropriate and useful for intended purposes. For teacher certification, the intended purpose is identifying individuals ready for professional practice, and the critical inference is that satisfactory pre-service performance indicates adequate professional teaching competence. This inference requires predictive validity evidence—empirical demonstration that pre-service performance correlates with independent measures of teaching effectiveness.

Messick identified multiple facets of validity, including content representativeness (does assessment content appropriately sample the construct domain?), substantive meaning (do assessment tasks actually engage the theoretical processes underlying the construct?), structural fidelity (do assessment scoring models reflect the theorised structure of the construct?), generalisability (do assessment results generalise across relevant contexts?), external relationships (do scores correlate as theoretically predicted with external criteria?), and consequential basis (are the social consequences of assessment use consistent with intended purposes?). For teacher education assessment, each facet raises important questions, but predictive validity, whether pre-service scores predict professional performance, provides the most direct evidence regarding whether certification decisions appropriately distinguish those ready for practice from those who are not.

Empirical Research on Pre-Service Performance and Teaching Effectiveness

The studies of the relationship between pre-service preparation predictors and post-teaching efficacy have yielded varied results in the context of different educational settings and differed significantly in terms of the rigour of methods. Other researchers have established small positive relationships between specific features of pre-service programs and their effectiveness in teaching in terms of student learning gains, observation ratings, or other metrics (Maclellan, 2001;, Fland, et.al., 2012; Wren, et.al., 2009; Asghar, 2012; Ferguson, 2011), Fook & Sidhu (2011). An example is the studies of teacher licensure scores, which have yielded small to moderate

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue V May 2026

correlations with student achievement gains, especially those tests that focus on content knowledge in the taught subject (Fook & Sidhu (2011); Akinbode & Dairo, (2011)). Nevertheless, the correlation coefficients have always been relatively small, with the aim of less than 10 per cent variance in teaching effectiveness.

Other studies have observed weak or negligible relationships between typical pre-service factors and teaching effectiveness. Research on the relationship between college GPA and teacher effectiveness has reported weak correlations, indicating that academic performance in college coursework is not a strong predictor of teacher effectiveness in classroom instruction. Likewise, studies about student teaching assessments have revealed that there is a low predictive validity to be applied, as the rating of supervisors during student teaching is weakly related to future teaching performance (Grossman, 2008; Adetayo, 2018; Dickson-Omogoye, 2025). Such results indicate that standard academic testing can capture skills that are not highly transferred to work life, such as theoretical knowledge that may not effectively translate into practical teaching abilities in the classroom.

Within the context of the Nigerian setting, in particular, there is scanty research done that explores the connections between pre-service training and professional competence. Although many studies have explored teacher education programs, most of them have assumed cross-sectional designs that study current practice, analysis of curriculum content, or stakeholder surveys, as opposed to longitudinal designs that would allow the testing of predictive relationships. The available longitudinal studies that exist have typically investigated relationships at program levels, as opposed to exploring the hypothesis that particular pre-service assessment scores are predictive of teaching competence.

The Nature of Teaching Competence

To understand why pre-service academic performance may not strongly predict teaching effectiveness, one must recognise that teaching competence is multidimensional and context-specific. Teaching involves not only knowledge of the subject matter but also pedagogical content knowledge that connects knowledge of the discipline with knowledge of how to teach a specific topic and common myths that students form. It needs classroom management skills that provide productive learning environments, diagnostic skills that detect when students are confused and modify teaching to meet the needs of the different students, and interpersonal skills that establish positive relationships with different students. Standard coursework that focuses on theoretical knowledge may not sufficiently practice or evaluate these applied skills (Umar, 2014).

Additionally, adaptive expertise, which is the capacity to act with flexibility in new circumstances and the needs of diverse students, as well as the unforeseen difficulties, is the dependent variable on teaching performance that cannot be developed in the pre-service training. Although theoretical knowledge can provide a foundation for adaptation, adaptive expertise primarily develops through long-term practice with feedback, making the growth of teaching competence a result of both professional experience and academic coursework (Idowu & Esere, 2009). In the event that this is the case, we would only anticipate modest, at best, correlations between academic achievement in pre-service courses and teaching effectiveness as measured by professional experience teachers have acquired.

METHODOLOGY

Research Design

The design used in this research was a concurrent explanatory mixed-methods design that combined quantitative correlation analysis, with qualitative observation and interview data. The sample population included NCE graduates who had taken ENG 212 (Applied English Linguistics) in the process of their pre-service training in Federal Colleges of Education in south-western Nigeria and were now teaching in primary or junior secondary schools in the area. By snowball sampling, 60 graduates were recruited who met the eligibility criteria: completed NCE programme in English Language Education with recorded grades in ENG 212 and were currently employed as a teacher with at least one year of professional experience and willing to be observed in a classroom and interviewed.

The 60 sample size was about 20 percent of the applicable cohort of the three Federal Colleges of Education that were part of the overall study. Although this purposive sample is not a probability sample, it was developed in such a way that the graduates of the entire spectrum of pre-service performance levels and of a variety of institutions were allowed to participate in the sample to increase the transferability of results.

Data Collection

Official institutional records were used to find pre-service performance data that records the final grade in ENG 212 of each participant based on both continuous assessment (40%) and examination (60%). The grades were between 40-100 with the institutional grading standards determining the levels of performance including Excellent (70-100), Good (60-69), Satisfactory (50-59), and Weak (40-49).

The Practicing Teacher Observation Rubric (r=0.87) was adapted and used as a measure of in-service teaching competence based on the structured classroom observation. Trained observers, who had gone through a standardized training to use the rubric, observed each graduate through two whole-class periods. Each dimension was independently rated by observers on a 5-point scale of Inadequate to Excellent, with descriptive anchors outlining each level. The average scores of overall teaching competence were obtained by summing the scores of dimensions and observation sessions, resulting in scores between 1 and 5. These continuous scores were further broken down into: Excellent (4.5-5.0), Good (3.5-4.49), Satisfactory (2.5-3.49), or Needs Improvement (below 2.5) to be used in chi-square analysis.

Qualitative data consisted of detailed field notes based on observation in the classrooms recording certain teaching behaviours, student reactions, and classroom dynamics as well as semi-structured interviews with 60 sampled graduates. Interviews investigated perceptions of adequacy of preparation reported by graduates, difficulties in professional practice, what helped them become effective or struggle in teaching practice, and links between their pre-service experiences and practice.

Data Analysis

The quantitative analysis involved Pearson Product-Moment Correlation analysis to test the linearity of the relationship between pre-service ENG 212 scores and in-service observation scores, both of which were continuous variables. Chi-square analyses were done to determine whether the categorical pre-service performance levels (Excellent, Good, Satisfactory, Weak) and categorical in-service performance levels were independent or had significant associations. All statistical tests were set at $\alpha=0.05$. Thematic analysis of qualitative data was done. Observation notes and interview transcripts were coded systematically to determine patterns in the teaching behaviours, common challenges or strengths, and factors that graduates reported as having been significant to their development as teachers. Special emphasis was placed on those cases with high divergence between pre-service and in-service performance to get insights into what could be behind some weak pre-service performers becoming good teachers and some strong pre-service performers becoming poor in practice.

FINDINGS OF THE STUDY

1. Correlation between Pre-Service and In-Service Performance

Table 1 : Pearson Product Moment Correlation table showing correlation between Observation of NCE graduates pre-service and in-service performances

Variables	N	Mean	Std. Deviation	R	Sig.
In-service Observation	51	78.1275	18.01717	.214	.132
Assessment Practices	51	69.0588	4.71768		

The Pearson product-moment correlation analysis in Table 1, was used to test the linearity of the relationship between the performance of the graduates during pre-service ENG 212 and their in-service teaching competence, which was assessed using classroom observation. Pre-service ENG 212 scores ranged from 42 to

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue V May 2026

88 (M=65.3, SD=12.4), while in-service observation scores ranged from 1.8 to 4.9 (M=3.4, SD=0.82 when converted to an equivalent scale, M=78.1, SD=18.0 on a 100-point scale).

Pre-service performance and in-service competence were correlated, r = 0.214, p = 0.132. The positive correlation means that as the pre-service scores were higher, the in-service scores were more likely to be higher, but the connection was weak and statistically insignificant at the 0.05 level. The square of the correlation coefficient (0.214) is approximately 0.046, which means that it was able to explain only 4.6% of the variance in professional teaching competence. In other words, the fact that the ENG 212 score of a graduate can tell us little about their future teaching performance – more than 95 per cent of the variance in teaching competence is unrelated to ENG 212 performance.

The fact that the p-value is not significant (0.132 > 0.05) does not mean that we can be certain that there is even a weak positive relationship in the general population; this uncertainty can be due to the fact that the relationships can be due to mere chance in this sample. Thus, the results cannot be used to prove the predictive validity hypothesis, according to which the pre-service performance of ENG 212 predicts professional teaching competence.

2. Chi-Square Analysis of Categorical Performance Levels

Table 2a: Grade * Categorised Score Crosstabulation

Count

		Categorised Score				**Total**
		Not satisfactory	**Needs Improvement**	**Good**	**Excellent**	
Grade	Distinction	1	1	0	2	4
	Upper Credit	5	5	5	2	17
	Credit	6	7	5	6	24
	Lower Credit	0	1	4	1	6
Total		12	14	14	11	51

The chi-square analysis in Table 2a, was performed to determine whether there was any significant association between pre-service performance categories such as 'Excellent', 'Good', 'Satisfactory' and 'Weak' and in-service performance categories such as 'Excellent', 'Good', 'Satisfactory' and 'Needs Improvement'. The number of graduates per category in pre-service was: Excellent (n=12), Good (n=21), Satisfactory (n=19), and Weak (n=8). In-service categories distribution was: Excellent (n=9), Good (n=24), Satisfactory (n=20), and Needs Improvement (n=7).

Table 2b: Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Approximate Significance
Pearson Chi-Square	9.321[a]	9	.408	
Likelihood Ratio	10.658	9	.300	
Linear-by-Linear Association	.651	1	.420	
Nominal by Nominal Phi	.428			.408
Nominal by Nominal Cramer's V	.247			.408
N of Valid Cases	51			

a. 12 cells (75.0%) have expected count less than 5. The minimum expected count is .86.

Table 2b shows Chi-square test results: χ^2=9.321, df=9, p=0.408. The insignificant value (p = 0.408 > 0.05) shows that there is no significant relationship between the pre-service and in-service performance types. This implies that the level of categorical performance of graduates in the pre-service training did not indicate the level of categorical performance of graduates in professional practice. There was no significant difference in the likelihood of excellent pre-service performers being excellent in-service performers compared to satisfactory or weak pre-service performers.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue V May 2026

Analysis of the cross-tabulation showed trends in agreement with this null result. Of the 12 graduates who scored 'Excellent' in ENG 212, only 3 (25%) of them got the rating of 'Excellent' in teaching, 6 (50%) got the rating of 'Good', 2 (16.7%) got the rating of 'Satisfactory' and 1 (8.3%) got the rating of 'Needs Improvement'. Out of the 8 graduates who did not score high on ENG 212, 2 (25%) of the students scored 'Excellent' in teaching, 3 (37.5%) of the students scored 'Good', 2 (25%) scored 'Satisfactory', and 1 (12.5%) scored 'Needs Improvement'. The similar distributions of the in-service categories irrespective of pre-service category depict the absence of a predictive relationship.

3. Qualitative Findings: Divergent Trajectories

Extensive classroom observations and interviews showed remarkable patterns of divergence between pre-service and in-service performance that can be utilised to explain the weak correlation observed quantitatively. Several graduates with low academic success in pre-service training, whose ENG 212 scores were in the 40s or in the low 50s, achieved very competent teaching practice, which was manifested by a number of strengths.

One of the teachers, (pre-service score: 47) dealt with 40 students in an amazing classroom setting with outstanding skills, keeping the place productive due to the variety of activities, strategic pacing, and the ability to manage positive behaviour. Her lesson on verb tenses involved a clear explanation with several examples; guided practice with instant feedback; practice with independent practice with continuous support; and formative checks with most of the students showing that they have mastered the concept. Interviewed on her pre-service experience, she said, 'I had problems with the written exams since I had trouble in English, but once I began teaching, I knew I could teach things in a manner the students could comprehend. I observe them keenly, and when they appear lost, I will attempt otherwise. The books did not teach me that – I learnt it by observing other teachers and by experimenting with my students.'

Another teacher, (pre-service score: 51) showed advanced pedagogical content knowledge, predicting typical student mistakes in sentence structure and setting up activities that would focus on correcting these mistakes. His classroom environment was encouraging and supportive with a high standard. He answered about his growth: 'I memorised during exams in college, but I didn't really understand the teaching process until practicum. And even then, I was simply doing the format we were instructed. The actual teaching that I got during my first two years of school, where I failed and had to determine what works. I could see what the courses wanted me to learn, but I learned by doing, not studying.

On the other hand, several graduates who had high pre-service performance (scores of 70s and 80s) had alarming weaknesses in practice. A teacher (pre-service score: 76) demonstrated poor classroom management where students were often off-task and disruptive, lessons were taught in a superficial way with no follow-up of comprehension, and there was a lack of student engagement. Students found her precise but overly technical explanations difficult to comprehend. During the interview, she acknowledged her struggles, stating, "While I excelled in all my courses and mastered the content, managing students is a completely different challenge than passing exams." I am aware of what to teach, and students are not cooperating as I expected. The classroom is disorganised, and I am not sure how to manage it."

The other teacher (pre-service score: 82) had good knowledge of the subject but had inflexible pedagogy, lecturing all the time with little student involvement and no formative assessment as the lesson progressed and could not change when students seemed lost. Students were not very engaged, and many students passively copied notes. The interview found out that I teach as I was taught in my courses – to present the information clearly and systematically. Students do not appear to learn as well as I did. Making it more interesting seems to require covering less content, which I want to avoid.

4. Factors Contributing to Divergence

Interview and observation data identified various factors that contribute to understanding why there is a gap in pre-service and in-service performance. To begin with, the competencies evaluated in ENG 212, which are mainly theoretical knowledge about linguistics, grammatical analysis, and language structure, are radically different to competencies needed to teach well. Academic achievements in language structure did not always

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue V May 2026

translate to success in the area of providing language to young learners, and those who had no success in linguistic theory were occasionally able to intuitively grasp how to teach language to young learners.

Secondly, classroom management was a key competency that was poorly covered during pre-service training and not evaluated during ENG 212. Several successful teachers listed classroom management as the most useful skill they had acquired by practice and not training, and several struggling teachers listed it as their greatest challenge. This implies that pre-service assessment can overlook or under-value competencies that are important to teaching success.

Also, the most useful part of pre-service preparation identified by the graduates was practicum, which was, however, too short (six weeks) and assessed according to the format of the lesson plans instead of teaching effectiveness. Some said they did well in practicum by following the format and using scheduled lessons to show observers, but that didn't help them make quick decisions in full-time teaching.

Furthermore, individual factors such as resilience, self-reflection, desire to seek assistance, and dedication to students were found to be more predictive of teaching effectiveness than academic ability. Successful teachers independent of pre-service performance explained actively pursuing mentorship, observing other teachers, trying out alternative methods, and persevering during challenges. Less effective teachers reported feeling lonely, confused about what to do to get better or blaming the student traits instead of analysing their practice.

DISCUSSION

1. Implications for Predictive Validity

The conclusion that there is weak and non-significant correlation between pre-service ENG 212 performance and in-service teaching competence brings up inherent issues regarding the predictive validity of the assessment practices in colleges of education. Predictive validity involves showing that the scores of an assessment are predictors of the criterion performance. In the case of teacher certification, professional teaching competence is the final factor – the factor that certification decisions are to predict. The results of this research indicate that the performance of ENG 212 is not effective in predicting this measure.

We can interpret this result through the lens of validity theory, which raises doubts about the legitimacy of ENG 212 performance as a factor in certification decisions. In the event that the pre-service grades do not predict teaching competence, what are they measuring and why should they be included in certification? The most philanthropic is that they gauge theoretical knowledge of linguistics which, though possibly useful as a general education, is not applied to practical teaching competence. The less philanthropic way to think about it is that they mostly assess the ability to take tests or academic skill at contextualised tasks that have little similarity with the real world of practice.

Both interpretations imply that current assessment practices do not have the predictive validity to warrant their use in high-stakes certification. Messick (1989) presents his validity paradigm to stress that validity is not just a technical issue of measurement but has social and ethical aspects – invalid tests result in incorrect decisions which are harmful. When the existing testing systems do not detect teaching ability, then they may examine people who are not good at teaching the job and miss out on people who would be good at teaching, which is detrimental to both the individual being tested and the children that such incompetent teachers will teach.

2. Systems Theory Perspective on the Disconnect

In terms of systems theory, the fact that pre-service performance (the result of the assessment) has a weak correlation with in-service competence (the result in the end) reflects severe dysfunction in the system. Professional competence and assessment in a well-operating teacher education system should be predictive of each other since the two should share the same underlying construct-teaching capability. A failure to establish such a relationship indicates that the system may be measuring irrelevant factors, developing inappropriate competencies, or both. (OECD, 2019)

Qualitative reports show conflicting paths where poor academic achievers emerge as good teachers, while bright academic achievers become bad teachers, indicating that existing teacher education programmes might be tailored to generate good students rather than good teachers. Academic assessment skills are acquired by memorising, theorising, and learning to take tests, but these skills do not transfer perfectly when it comes to classroom instruction, which requires practical skills, adaptive expertise, and interpersonal competencies.

There are implications of feedback loops with regard to this system dysfunction. Unless assessment is used to measure teaching competence, it cannot provide valid feedback to inform improvement in the programme of the guide. The assessment system used by colleges fails to measure teaching skills, and as such, colleges cannot determine whether their programmes are producing teaching skills. This discontinuous feedback loop makes the system unable to correct itself, and the discrepancy between training and the practice continues, leading to a persistent gap in teacher effectiveness and student outcomes (Wren, et.al., 2019; Asghar, 2012; Ferguson, 2011; Fook & Sidhu, 2011).

3. The Theory-Practice Gap in Teacher Education

The results illuminate the persistent challenges in teacher training worldwide, particularly regarding the relationship between theoretical knowledge and practical competence. Education programmes in teacher education have always had difficulties in marrying coursework based on theoretical knowledge with practical experiences based on the development of teaching skills. The results of this paper indicate that the pendulum might have swung too far in the Nigerian context in favour of theoretical knowledge assessment, to the detriment of practical competency assessment.

Qualitative data indicated that successful teachers would frequently explain how they learnt to teach: by doing it, trial and error, observing other teachers and thinking about what went well and what did not. This implies that competence in teaching is built up by practice over an extended period with feedback and not by coursework. On the one hand, theoretical knowledge offers the base; on the other hand, there is a need to transfer the knowledge into a working practice which necessitates a chance to apply, experiment, fail, get feedback, and revise approaches, and, perhaps, brief practicum experiences may not offer such opportunities.

This approach does not only have assessment curriculum implications. Assuming that teaching competence is largely developed as a result of practice, teacher training ought to offer more extensive, closely observed, and systematically evaluated practice experiences rather than offering theoretical instruction. The existing paradigm of two-three years of coursework and a short practicum could be essentially ineffective in the emergence of teaching competence, which would justify why pre-service academic achievement does not forecast professional effectiveness.

4. Rethinking What Should Be Assessed

The lack of correlation between pre-service and in-service performance poses important questions concerning what is to be measured in teacher training. When theoretical knowledge on traditional academic assessments fails to predict teaching competence, what is to be assessed? Anecdotal observation and interview results indicate that there are a number of competencies that were identified in effective and struggling teachers but possibly not well-assessed in pre-service training.

Classroom management skills were identified as very important and seemingly underestimated. Teachers who were effective provided effective learning environments by being organised, managing behaviour in ways that helped to create an effective environment, utilising time well, and developing warm yet well-structured classroom climates. These skills seem teachable and testable but can be overlooked in course work that focuses on content rather than pedagogy and can be under-tested on a brief practicum.

Pedagogical content knowledge; knowledge of how to make content available to students, of what challenges can be expected, of how learning activities should be designed, is a characteristic of effective teachers but may not be sufficiently developed or tested in courses that focus on either pure content or generic pedagogy unless it is sufficiently integrated. The ability to diagnose when the students are lost in their learning and make

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue V May 2026

adjustments in their instruction, answer unanticipated questions or situations, and alter plans as formative assessment seems to be the core of teaching, yet it is hard to measure in terms of practicum-planned lesson demonstrations. This evidence implies that further time-prolonged evaluation is needed in real-world situations where adaptive decisions are being made in real life.

Professional dispositions such as commitment to students, readiness to seek assistance and be a lifelong learner, ability to reflect on practice and to analyse it and improve practice appeared to be significant but are not often formally measured. Current assessment might implicitly presuppose that anyone who is doing coursework has the right dispositions, but there is evidence that this presumption should be tested.

5. Comparative Perspective

The results of this study are consistent with other larger global studies that indicate small at best correlations between pre-service predictors and teacher performance. Pivotal research in different countries has always determined that conventional pre-service predictors, such as GPA, test results, and coursework grades, account for fairly trivial amounts of variation in teaching efficiency, as indicated by student learning gains or observational ratings. This implies that the issue is not particular to the Nigerian Colleges of education but a wider issue concerning teacher education in terms of the lack of coherence between practice and preparation.

Nonetheless, the specifically low correlation in this study (r=0.214) in comparison with a moderate positive correlation in other settings (usually r=0.20 to 0.40) indicates that the misalignment could be even more extreme in Nigerian colleges of education than in some other settings. This may indicate a few things, such as theoretical knowledge that may have been overemphasised at the cost of practical training; poorly developed or short practicum experiences; poorly designed assessment tools that may not be well-adjusted to the realities of teaching; and no performance-based assessment that concentrates on demonstrated teaching competencies.

CONCLUSION

This research offers empirical data on one of the key validity questions of teacher education: Does pre-service performance correlate with professional teaching competence? It is evident in the evidence that in the situation under study, the performance of NCE graduates in ENG 212 in colleges of education in south-western Nigeria, the performance in pre-service fails to predict the in-service competence. The non-significant (p=0.132) weak correlation (r=0.214) indicates that pre-service grades only accounted for about 4.6% of variance in teaching effectiveness, and more than 95 per cent of the variation in professional competence is attributed to other factors besides pre-service academic performance.

This discovery has far-reaching consequences. Assessments in pre-service training cannot serve their basic role in a certification system when they are not useful in trying to predict teaching competence. Empirical justification of certification decisions relying heavily on tests which have low predictive validity is impossible and could result in the certification of those who will not perform well in practice and the disqualification of those who will perform well. This is not only a technical issue about measurement but also a systemic problem that has implications about the quality of education in schools that employ graduates of NCE.

Qualitative evidence assists in understanding the reasons behind the deviation in pre-service and in-service performance. The competencies that are evaluated academically, mainly the theoretical knowledge of linguistics, are very different from those that are needed in the work of an effective teacher, such as classroom management, pedagogical content knowledge, adaptive instructional decision-making and the ability to work with various learners. Sometimes, graduates with high academic achievement were deficient of real-life teaching skills, whereas those with low academic achievement sometimes showed a high level of sophisticated teaching competence, achieved through experience and reflection. This implies that existing teacher education programmes can be designed in such a way that they only result in good students but not good teachers.

In terms of systems theory, the lack of connection between pre-service performance and professional competence points to the failure of feedback loops and the dysfunction of the entire system. The system can produce performance information by evaluation, yet this information is not a valid sign of the criterion of

interest, teaching competence, and so it cannot allow valid inferences and significant quality assurance. Teacher education can only have the predictive validity that certification systems demand by fundamental reconceptualisation of what is measured and how.

RECOMMENDATIONS

The reform of the validity failures reported in this study necessitates an extensive reform at various levels. These recommendations are divided into short-term measures that can be taken relatively soon and structural reforms that are long-term and demand more fundamental changes.

Immediate Actions:

1. Adopt Performance-Based Assessment: Substitute or add to written exams with performance-based assessment, which involves showing teaching skills in real-life situations. Consider video-taped instructional demonstrations scored to rubrics, instructional performance portfolios, and assessed microteaching lessons where pre-service teachers instruct peers with structured feedback.

2. Extend and Strengthen Practicum Assessment: Add to the existing six-week practicum at least one full semester of daily teaching duties. Facilitate holistic assessment rubrics of practicum based on professional teaching standards; train supervising teachers and college supervisors in systematic observation and rating; and mandate multiple observations over the course of the practicum in order to measure teaching competence holistically and not just on single demonstration lessons.

3. Build Explicit Teaching Competency Frameworks: Based on NCCE standards and research on effective teaching, create explicit competency frameworks of what is known, what skills and what dispositions are needed to be effective teachers. Make sure that all programme elements such as courses, practicum and assessments explicitly cover particular competencies in the framework. Make outlines clear to the learners to know what they have to prove.

4. Adopt Cumulative Teaching Performance Assessment: Instead of practicum being an independent component that is assessed independently, adopt performance assessment across the programme where continued development of increasingly advanced teaching competencies is demonstrated. This may start with microteaching in introductory courses, advance to field experiences in schools and finally culminate with the comprehensive practicum, which is a demonstration of full competence in teaching.

Long-Term Structural Reforms:

1. Carry out Longitudinal Validation Studies: Develop systematised mechanisms of following up graduates into practice and studies of connections among different pre-service pointers and teaching efficiency gauged in a variety of methods, including classroom observations, student learning gains, and principal assessments. Apply results to further hone assessment practices, determining which of the pre-service assessments are in fact predictive of professional competence and revising or abolishing those that are not.

2. Redesign Teacher Education Curriculum: shift away from heavy focus on theoretical content to more balanced focus on theory and practice in the programme. Integrate practical experience in content courses; demand more and earlier field experiences; and structure curriculum around acquisition of particular teaching skills instead of academic content.

REFERENCES

1. Adeleke, J.O. and Dickson-Omogoye, B.T., (2003). An analysis of effects of test misconduct on students' perception of assessment quality in Nigeria federal colleges of education. The African Journal of Behavioural and Scale Development Research. Doi: 10.58579/ajb-sdr/5.2.2023.18
2. Adetayo, J. O. (2018). Teacher quality and instructional effectiveness in Nigerian secondary schools. *International Journal of Educational Research Review, 3*(2), 45–52.

ISSN No. 2321-2705 | DOI: 10.51244/IJRSI |Volume XIII Issue V May 2026

3. Akinbode, J. O., & Dairo, M. D. (2011). Influence of teacher qualification on students' academic performance in Nigeria. *Journal of Educational and Social Research, 1*(2), 45–52.
4. Anderson, S. K., Ozsezer-Kurnuca, S., & Jain, P. (2024). Judging student teacher effectiveness: A systematic review of literature. *British Journal of Educational Studies, 72*(5), 553–585.
5. Asghar, M. (2012). Teacher preparation and student achievement: A review of evidence. *Journal of Education and Practice, 3*(12), 45–50.
6. Darling-Hammond, L. (2006). Powerful teacher education: Lessons from exemplary programs. Jossey-Bass.
7. Darling-Hammond, L., Flook, L., Cook-Harvey, C. Barron, B. and Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science, 24(2), 97-140.*
8. Dickson-Omogoye, B.T. and Kehinde-Dada, O.V. (2026). Enhancing error-visibility of assessment practices for sustainable educational reform: a monitoring and evaluation perspective. *International journal of research and innovation in social science (IJRISS)*, 10(26), 1774 -1784.
9. Campos, F. C., Freitas, L. C., & Ribeiro, V. M. (2017). Assessment validity in educational contexts: A systematic review. *Studies in Educational Evaluation, 54*, 1–10.
10. Easton, D. (1965). A framework for political analysis. Prentice-Hall.
11. Ferguson, R. F. (2011). Teacher effectiveness and student achievement. *Educational Evaluation and Policy Analysis, 33*(2), 123–145.
12. Fook, C. Y., & Sidhu, G. K. (2011). School-based assessment among ESL teachers. *Journal of Education and Learning, 5*(2), 93–98.
13. Grossman, P. (2008). Responding to our critics: From crisis to opportunity in research on teacher education. *Journal of Teacher Education, 59*(1), 10–23.
14. Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, re-imagining teacher education. *Teachers and Teaching, 15*(2), 273–289.
15. Idowu, A. I., & Esere, M. O. (2009). Assessment in Nigerian schools: A counselor's viewpoint. *Edo Journal of Counselling, 2*(1), 17–27.
16. Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), Educational measurement (4th ed., pp. 17 64). American Council on Education and Praeger.
17. Maclellan, E. (2001). Assessment for learning: The differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education, 26*(4), 307–318.
18. Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). Macmillan.
19. National Commission for Colleges of Education. (2012). Minimum standards for NCE teachers. NCCE.
20. OECD. (2019). A flying start: Improving initial teacher preparation systems. OECD publishing.
21. Roloff, J., Klusmann, U., Lüdtke, O., & Trautwein, U. (2020). The predictive validity of teachers' abilities on instructional quality. *AERA Open, 6*(1).
22. Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1), 1-22.
23. Singh, K. (2014). *Test measurement and research methods in behavioural sciences*. Bharati Bhawan.
24. Umar, A. (2014). Teacher effectiveness and classroom performance in Nigeria. *Journal of Educational Foundations, 5*(1), 89–102.
25. Wren, D. G., et al. (2009). The relationship between teacher preparation and student achievement. *Educational Research Quarterly, 33*(1), 3–18.
26. Von Bertalanffy, L. (1956). General system theory. General Systems, 1(1), 1-10.
27. Zeichner, K., & Conklin, H. G. (2008). Teacher education programs as sites for teacher preparation. In M. Cochran-Smith, S. Feiman-Nemser, & D. J. McIntyre (Eds.), Handbook of research on teacher education (3rd ed., pp. 269-289). Routledge.