

## RESEARCH ARTICLE

# Reliability of the Philippine Nurse Licensure Examination (PNLE)

James Montegrigo, PhD, RN, CMSRN<sup>1</sup>, Geraldine Ferreras, PhD(c), MSN, RN<sup>2</sup>  
and Kelly Powers, PhD, RN, CNE, FNAP<sup>3</sup>

## Abstract

**Aim:** To determine the reliability of the Philippine nurse licensure examination (PNLE) results from 2010 to 2019.

**Background:** The PNLE first-time pass rate (FTPR) is a major criterion used to evaluate nursing program quality. There is wide variability in PNLE results ranging from 0% to 100% across nursing schools. There are no published studies to investigate the reliability of the PNLE.

**Methods:** A retrospective correlational research design was used to determine the reliability of PNLE. Secondary data analysis of 517 nursing schools' PNLE FTPR results was conducted. Descriptive statistics described trends in PNLE results over 10 years. Spearman correlation was used to determine PNLE reliability.

**Findings:** The PNLE FTPR increased from 2010 to 2019. More than half of Philippine nursing schools have low PNLE performance. Combined annual PNLE results, that is, adding May/June and November/December results, provided a more accurate measure of reliability.

**Conclusion:** The PNLE is a reliable measurement tool to evaluate nursing program outcomes. Regulatory bodies should use caution in using PNLE results as a criterion in assessing nursing program quality.

**Keywords:** *nursing education, nurse licensure examination, Philippines, board exam*

## Introduction

The first-time pass rate (FTPR) in the nurse licensure examination (NLE) is a major criterion used in evaluating nursing program quality (Commission on Higher Education [CHED], 2017; National Council State Boards of Nursing [NCSBN], 2019). The NLE outcome is a multifactorial phenomenon that involves an interplay of student characteristics, programmatic factors, institutional variables, and the examination itself (Bautista et al., 2018; Montegrigo, 2020; Rosales et al., 2014; Waltz et al., 2017). Regulatory agencies use the Philippine nurse licensure examination (PNLE) FTPR as a benchmark in monitoring nursing program quality and in imposing penalties on low-performing nursing schools (CHED, 2017). Earlier studies have expressed concerns about the quality of nursing education in the Philippines, based on PNLE results from 2006 to 2018 (Bautista et al., 2018; Montegrigo, 2020; Rosales et al., 2014).

Licensure examinations are high-stakes examinations (Foreman, 2017; McCoy, 2019; Waltz et al., 2017) because they are used to determine the competency of professionals entering nursing practice and to make decisions about nursing program quality. Thus, licensure examinations should be psychometrically sound and legally defensible (NCSBN, n.d.). High-stakes examinations must have strong reliability (Foreman, 2019; McCoy, 2019; Waltz et al., 2017). To date, there are no published studies about the reliability of the PNLE. Therefore, the consistency of the PNLE in measuring nursing graduates' competencies and evaluating nursing program quality is unknown. This was the first study to analyze the reliability of the PNLE over 10 years (2010 to 2019).

This study described the PNLE FTPR of 517 nursing schools and determined the stability of these results over 10 years. The

<sup>1</sup> Assistant Professor, School of Nursing, The University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, North Carolina, USA Correspondence email: jmontegr@unc.edu

<sup>2</sup> Faculty Member, School of Health and Natural Sciences, Saint Mary's University, Bayombong, Nueva Vizcaya, Philippines

<sup>3</sup> Associate Professor, School of Nursing, The University of North Carolina at Charlotte, 9201 University City Blvd., Charlotte, North Carolina, USA

study is significant for several reasons. The results of this study can provide input to stakeholders such as nursing schools, CHED, and the Board of Nursing (BON), about the reliability of PNLE over time. Moreover, it provides evidence-based information to guide policy development and implementation, particularly in providing sanctions to graduating nursing students and nursing schools.

## Review of Literature

### Test Reliability

Reliability testing is based on two measurement frameworks, norm-referenced and criterion-referenced frameworks. The PNLE uses a criterion-referenced framework because to pass the PNLE, an examinee must achieve an average of 75% and a grade of 60% in any of the five parts of the examination (CHED, 2017). In educational research, test reliability refers to the “consistency with which a tool or method assigns scores to subjects” (Waltz et al., 2017, p. 7). This means that if the PNLE does not consistently measure the graduates' competencies or nursing program quality, its usefulness is questionable. A critical aspect of reliability, therefore, is the stability of the measure (the PNLE) over time.

The reliability of the PNLE may be affected by scoring, exam characteristics, individual attributes, and the environment in which the measure is administered, thus causing wide variability in test outcomes (Waltz et al., 2017). For the PNLE to be reliable, FTPR results over time from parallel or similar versions of the PNLE should demonstrate a high correlation (Furr & Bacharach, 2014). An instrument that is not reliable has questionable validity (Foreman, 2017; McCoy, 2019; Waltz et al., 2017).

As evidence for reliability accumulates over time, it is imperative to continuously and progressively describe the reliability results of a measurement tool (Waltz et al., 2017). While there is a lack of research on the reliability of the PNLE, Foreman (2017, 2019) argued that the application of NLE results to evaluate the quality of nursing programs needs to be reviewed because of the wide variability in FTPR. This study provided preliminary evidence for this purpose.

### The Philippine Nurse Licensure Examination

Licensure examinations serve different purposes: to ensure public safety, maintain institutional and program reputation, and serve as a requirement for accreditation, thus making these exams high-stakes for institutions, faculty members, and students (National League for Nursing [NLN], 2012). In the Philippines, the PNLE is administered nationwide over two days twice a year. The first PNLE is administered in May/June and the

second PNLE is administered in November/December. The first exam day covers Nursing Practice (NP) I (Community Health Nursing), NP II (Care of Healthy/At Risk Mother and Child), and NP III (Care of Clients with Physiologic and Psychosocial Alterations-Part A). The second day covers NP IV and NP V (Care of Clients with Physiologic and Psychosocial Alterations-Parts B and C, respectively) (PRC-BON, 2019). The PNLE test blueprint has evolved over the years to reflect the changes in the Philippine nursing curriculum. It covers foundational biological sciences and clinical nursing subjects in the curriculum to ensure the representativeness of domains (PRC-BON, 2019). A test must cover the domains in the curriculum (Waltz et al., 2017). The PNLE questions, developed by members of the PRC-BON based on their areas of clinical specialization, consist of 500 multiple-choice questions (MCQ) divided into five parts. MCQ is commonly used in licensure examinations because it is the most objective, most reliable, and provides the most functionality in measuring all types of knowledge (Waltz et al., 2017).

Over the years, the PNLE results have exhibited wide variations and inconsistencies. Schools designated as centers of excellence in nursing education that consistently produced 95% to 100% FTPR in the PNLE, have experienced 0% to 50% FTPR. Conversely, nursing schools with consistently low FTPR occasionally achieved high FTPR. The wide variability of PNLE results points to the questionability of the PNLE as a reliable measure for high-stake decisions made by regulatory and accrediting agencies.

Previous PNLE studies reported a national FTPR average of 74.6% (Montegrigo, 2019) and an overall national pass rate (FTPR and repeat test taker pass rate) ranging from 49% in 2011 to 83.7% in 2021, suggesting a very wide variability in PNLE performance (Bautista et al., 2018; BON, 2021; Rosales et al., 2014). It can be inferred that Philippine nursing schools have high graduation rates, but have low PNLE FTPR. This creates concerns about the knowledge gap among nursing graduates or doubts about the reliability of the PNLE in measuring graduates' competencies.

### The Philippine Nursing Curriculum

The nursing curriculum in the Philippines is a four-year baccalaureate program that aims to produce generalist nurses. The curriculum has undergone several revisions to reflect the required competencies and address the needs of an ever-evolving healthcare system. Changes in the nursing curriculum may have influenced the PNLE itself and the performance of graduates in the examination. Incorporation of review classes during the senior year is a controversial issue in Philippine nursing schools. In an attempt to achieve a high

PNLE FTPR, comprehensive review programs that are either taught by nursing faculty or outsourced to commercial review centers, are conducted during the last semester of the nursing program or an in-house review is provided after graduation. Masselink and Lee (2010) claimed that at one point, almost 90 percent of nursing schools in the Philippines had business associations with nursing review centers, which caused commercialization and compromised the nursing education quality.

### The PNLE as a Measure of Quality Nursing Education

The PNLE test plan outlines the concepts covered in the examination and shows clear alignment of the subjects in the BSN curriculum. NP I covers Foundations of Nursing Practice, Nursing Research, and Leadership and Management while NP II measures competencies in Maternal-Newborn Nursing and Community Health Nursing. NP III, NP IV and NP V provide a comprehensive measurement of competencies in Medical-Surgical Nursing and Psychiatric Nursing (BON, 2017). The alignment of subjects in the BSN curriculum and the PNLE test plan is one way of ensuring the representativeness of the domains tested in the licensure examination (McCoy, 2019; Waltz et al., 2017). This domain-sampling model constitutes an essential indicator of reliability (Waltz et al., 2017). As a high-stakes test, government regulatory agencies impose penalties for nursing schools that have a low PNLE FTPR. Schools that have FTPR below 30 percent within three years are recommended for closure, while those that achieve 30 to below 50 percent are given a warning (CHED, 2017).

Foreman (2017) conducted a reliability study of the National Council Licensure Examination-Registered Nurses (NCLEX-RN) based on test results of 2,157 nursing programs across the United States from 2010 to 2016, which became the main basis for the methods and data analysis in this study. In this study, substantial variations in NCLEX pass rates were found, suggesting problems with the exams' validity and reliability. Foreman (2017) further argued that licensure examination rates can create erroneous conclusions regarding the quality of nursing education because the FTPR is a single fixed rating. The FTPR has a wide variation that can be attributed to systematic and random errors due to differences in student training, student characteristics, familiarity with the test, changes in the test plan, and regular changes among the board of examiners (Archer et al., 2016; Foreman, 2017; Montegrigo, 2020; NCSBN, 2019). Caution should be taken in using the PNLE as an evaluative measure since random and systematic errors may have occurred during exam administration. Single-test administration is not sufficient to determine reliability when it is administered on different occasions or with different participants (Waltz et al., 2017).

Currently, there are no published studies on the reliability of the PNLE. Therefore, the reliability of the PNLE as an evaluative measurement of the quality of nursing programs is largely unknown. This study on the reliability of the PNLE was the first research to be conducted in the Philippines and it addressed this very relevant gap in the literature.

## Methods

### Research Questions

This study aimed to assess the reliability of the PNLE. Two research questions were formulated to achieve this aim:

1. What is the trend and level of performance of Philippine nursing schools based on PNLE FTPR from 2010 to 2019?
2. Is the PNLE FTPR a reliable means to measure nursing graduates' competencies, determine the quality of nursing education, and provide a basis for imposing regulatory sanctions on schools with low PNLE performance?

### Research Design

A retrospective correlational research design was used to determine the reliability of PNLE. In educational measurement, correlational research design is an appropriate approach to establish the reliability of measures (Waltz et al., 2017). Correlation coefficients of annual and semi-annual PNLE results were calculated to determine the reliability of the PNLE.

### Sample

PNLE FTPR results of 517 nursing schools from 2010 to 2019 were analyzed. However, with the significant reduction in nursing enrollment across the country over the years, there were only more than 300 nursing schools in operation at the time this study was conducted. For correlational studies, a priori G\*power analysis required 84 schools as units of analysis to achieve 0.80 power, 0.30 effect size, and 0.05 level of significance. A 0.30 effect size is commonly used to establish a moderate effect size between the research variables while a power of 0.80 is generally accepted a priori guideline to predict an 80% chance of achieving statistical significance (Tabachnick & Fidell, 2013). It was critical to establish these values at the conceptualization stage of the study because failure to do so may result in an inability to find significant effects in the study (Tabachnick & Fidell, 2013).

### Data Collection

Research data were collected from publicly available PNLE results from the PRC-BON website. Due to the COVID-19

pandemic, licensure examinations throughout the country were canceled in 2020 and early 2021. We collected the number of first-time applicants, first-time passers, and FTTPR of May/June and November/December PNLE for each year from 2010 to 2019. Manual data entry was done on an Excel spreadsheet. In order to ensure the accuracy of data entry, we reviewed and double-checked each other's collected data. To ensure consistency and accuracy in the units of analysis, we merged data from nursing schools that changed their names or had different versions of their school names.

### Data Analysis

Descriptive and inferential statistics using SPSS version 27 were used to analyze the data. Descriptive statistics such as frequency, percentage, and mean described the 10-year PNLE results. To determine the stability of PNLE over time, inferential statistics was used to determine the relationship between variables or the PNLE results. The non-parametric Spearman correlation was used due to violations in normality (Shapiro-Wilk test, Q-Q plot, and box plots) and linearity. Correlations are expressed from .00 (no correlation) to 1.00 (perfect correlation).

To guide our data interpretation, correlation coefficients below 0.40 were considered weak correlations, 0.40 to 0.69 as moderate correlations, 0.70 to 0.89 as strong correlations, and 0.90 and above as very strong correlations (Schober et al., 2018). Correlations were converted to percentages by squaring the coefficient values to explain the percentage of explained variance between two PNLE results (Waltz et al., 2017).

Several essential considerations were taken into account before data analysis. First, we differentiated zero percent FTTPR that is due to the actual failure rate versus zero percent that is due to no examinees, since the official PRC results reported them similarly as zero percent FTTPR. This is critical to differentiate because it may pose significant methodological and analytical issues. Second, schools without any examinees in a particular PNLE year were left blank and were considered as data that are "missing at random" (MAR), which poses less serious problems in data analysis (Tabachnick & Fidell, 2013). Third, the overall PNLE FTTPR in each year was computed by determining the total number of examinees and the total number of those who passed the first time from the two PNLE exams (May/June and November/December). This was done to eliminate bias related to the practice of selecting high-performing students to take the first PNLE in May/June, and then the lower-performing students taking the second PNLE in November/December. Some schools prefer to take the second PNLE schedule to give enough time for the students to prepare for the exam. To reduce bias and measurement error, we computed the overall annual FTTPR of two PNLEs, rather than getting the mean of the two PNLE exams every year. Fourth, in computing for the 10-year FTTPR, the

overall total number of first-time examinees and first-time PNLE passers was determined.

To determine the level of performance of nursing schools, the following categories were used: high, average, low, very low, and poor performance. The "high performance" category was based on the BON cut-off passing rate (80% FTTPR) when determining top-performing nursing schools. CMO 15, s. 2017, used the following FTTPR ranges in determining sanctions to nursing schools: 45% - 55% (warning), 31% - 45% (probation), and below 30% (phase-out). We adopted this classification to describe low, very low, and poor performance. PNLE FTTPR of schools were classified as "average" if their FTTPR is 56% to 79%, "low" for FTTPR of 46% to 55%, "very low" 31% to 55%, and "poor" for FTTPR of 30% and below (CHED, 2017).

Finally, to determine the influence of high variability of PNLE rates and in an attempt to demonstrate the true FTTPR, the 95% confidence interval (CI) was determined. This was necessary to reflect that the true FTTPR lies within the CI because these values provide "a range of values of a sample statistic that is likely to contain the true population value" (Keller & Kelvin, 2013, p. 87).

### Ethical Consideration

Institutional review board approval was not obtained since the study involved data collection from publicly available sources. Additionally, aggregate findings were reported, thus, protecting privacy and confidentiality of information of individual schools of nursing.

## Findings

### Trend and Level of PNLE Performance of Nursing Schools

Table 1 presents the PNLE FTTPR according to the year and month of examination. The 10-year FTTPR is 62.1% (n=202,162). There is a continuous increase in PNLE FTTPR within the 10 years, from 54.3% (n=49,353) in 2010 to 85.2% (8,685) in 2019. A significant increase in FTTPR from 2015 (76.1%, n=10,166) to 2019 (85.2%, 8,685) was observed. Generally, there is a higher number of first-time examinees (n=178,437) and FTTPR (65.9%) during the May/June PNLE compared to November/December PNLE first-time examinees (n=146,884) and FTTPR (57.6%). The PNLE in the year 2019 had the highest FTTPR (85.2%, n=8,685) while PNLE in 2010 had the lowest FTTPR (54.9%, n=49,353). A 90.3% decrease in the number of first-time examinees was observed from 2010 (n=90,942) to 2016 (n=8,821) and a 25% increase in the number of first-time examinees was seen from 2017 (n=8,149) to 2019 (n=10,188).

Table 1. PNLE First-time Pass Rates from 2010 to 2019

Year and Month of PNLE	Number of first-time examinees	Number of first-time passers	Mean FTPR
<b>2010</b>	90,942	49,353	54.3%
May/June	52,054	30,391	58.4 %
Nov./December	38,888	18,962	48.8 %
<b>2011</b>	77,021	45,591	59.2 %
May/June	44,400	29,514	66.5 %
Nov./December	32,621	16,077	49.3 %
<b>2012</b>	53,318	33,016	61.9 %
May/June	31,715	21,538	67.9 %
Nov./December	21,603	11,478	53.1 %
<b>2013</b>	31,630	19,152	60.5%
May/June	18,117	12,079	66.7 %
Nov./December	13,513	7,073	52.3 %
<b>2014</b>	23,262	16,428	70.6 %
May/June	11,886	7,894	66.4 %
Nov./December	11,376	8,534	75.0 %
<b>2015</b>	13,357	10,166	76.1 %
May/June	6,236	4,920	78.9 %
Nov./December	7,121	5,246	73.7 %
<b>2016</b>	8,821	6,800	77.1 %
May/June	4,082	3,153	77.2 %
Nov./December	4,739	3,647	77.0 %
<b>2017</b>	8,149	6,309	77.4 %
May/June	3,320	2,436	73.4 %
Nov./December	4,829	3,873	80.2 %
<b>2018</b>	8,623	6,662	77.3 %
May/June	3,254	2,701	83.0 %
Nov./December	5,369	3,961	73.8 %
<b>2019</b>	10,188	8,685	85.2 %
May/June	3,373	2,909	86.2 %
Nov./December	6,815	5,776	84.8 %
<b>Total</b>	325,321	202,162	62.1 %
May/June	178,437	117,535	65.9 %
Nov./December	146,884	84,627	57.6 %

Table 2 presents the classification of nursing schools based on PNLE FTPR within the 10-year study period. Less than half (45.7%, n=236) of the country's nursing schools had "average" to "high" performance, with 14.9% (n=77) of nursing schools in the "high performance" category. More than half (54.3%) of nursing schools were classified as "low" (14.5%, n=75), "very low" (21.1%, n=109), and "poor" performance (18.8%, n=97).

**Reliability of PNLE FTPR**

Table 3 exhibits the correlation coefficients of annual PNLE FTPR from 2010 to 2019. All PNLE FTPR are statistically significantly correlated at  $p=.01$ , ranging from moderate ( $r_s=.517$

to  $r_s=.678$ ) to strong correlations ( $r_s=.703$  to  $r=.817$ ). PNLE results from 2010 to 2014 have significantly stronger correlations ( $r_s=.703$  to  $.817$ ) compared to PNLE from 2015 to 2019 ( $r_s=.517$  to  $.678$ ). These statistically significant moderate to strong correlations are evidences of the high reliability of the PNLE results.

Table 4 shows the correlation matrix of each PNLE result within 10 years. While all PNLE correlation coefficients demonstrated statistical significance ( $p=.01$ ), the magnitude or strength of the correlations were weaker ( $r_s=.26$  to  $.63$ ) when PNLE results were analyzed individually, compared to the correlation

**Table 2.** Classification of Nursing Schools Based on 10-year PNLE FTPR (2010-2019)

	Frequency	Percentage
High performance (80%-100%)	77	14.9 %
Average performance (56%-79.9%)	159	30.8 %
Low performance (46%-55.9%)	75	14.5 %
Very low performance (31%-45.9%)	109	21.1%
Poor performance (0%-30.9%)	97	18.8%
<b>Total</b>	<b>517</b>	<b>100.0 %</b>

**Table 3.** Correlation Matrix of Nursing Schools' Annual PNLE FTPR from 2010 to 2019

	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019
<b>2010</b>	-	.809*	.800*	.758*	.703*	.650*	.612*	.616*	.572*	.609*
<b>2011</b>	-	-	.817*	.771*	.703*	.678*	.645*	.621*	.555*	.638*
<b>2012</b>	-	-	-	.767*	.712*	.650*	.618*	.623*	.565*	.622*
<b>2013</b>	-	-	-	-	.710*	.662*	.647*	.645*	.598*	.635*
<b>2014</b>	-	-	-	-	-	.644*	.613*	.541*	.517*	.587*
<b>2015</b>	-	-	-	-	-	-	.629*	.548*	.589*	.562*
<b>2016</b>	-	-	-	-	-	-	-	.546*	.537*	.631*
<b>2017</b>	-	-	-	-	-	-	-	-	.518*	.541*
<b>2018</b>	-	-	-	-	-	-	-	-	-	.588*
<b>2019</b>	-	-	-	-	-	-	-	-	-	-

coefficients when the PNLE results were analyzed cumulatively on an annual basis ( $r_s = .517$  to  $.817$ ). With statistically significant correlations, the findings consistently suggest that the PNLE remained relatively stable within the 10 years, indicating that the PNLE is a reliable measurement tool.

Table 5 presents examples of 95% confidence intervals (CI) and percentage of explained variances ( $R^2$ ) from randomly selected PNLE results. The sample 95% CI reflects correlation coefficients of randomly selected annual PNLE results, June-to-June PNLE results, and June-to-December PNLE results. The research findings showed that all correlation coefficients were

statistically significant ( $p = .01$ ) within the 95% CI. This means that the true PNLE results fall within the range that is a reflection of the true population value (Keller & Kelvin, 2013), indicating strong reliability of the PNLE.

On the other hand, the  $R^2$  values indicate the percentage of explained variance between two randomly selected PNLEs. For example, 65.4% of the variance in PNLE results from 2010 to 2019 can be explained by the 2010 and 2011 PNLE FTPR ( $r_s = .809$ ) and the remaining 34.6% of the variance in PNLE results can be explained by PNLE FTPR from years other than 2010 and 2011.

**Table 4.** Correlation Matrix of Each PNLE FTPR Result from 2010 to 2019

	J 10	D 10	J 11	D 11	J 12	D 12	J 13	D 13	J 14	D 14	J 15	D 15	J 16	D 16	J 17	D 17	J 18	D 18	J 19	D 19
J10	-	.51	.62	.42	.63	.47	.59	.40	.50	.47	.40	.41	.44	.33	.41	.44	.36	.43	.50	.42
D10	-	-	.53	.66	.53	.58	.46	.57	.41	.56	.42	.51	.45	.47	.39	.50	.34	.49	.41	.45
J11	-	-	-	.49	.63	.49	.62	.47	.55	.51	.50	.47	.49	.38	.44	.42	.40	.43	.39	.47
D11	-	-	-	-	.50	.61	.43	.57	.45	.54	.39	.48	.44	.46	.35	.48	.37	.39	.47	.42
J12	-	-	-	-	-	.51	.67	.42	.52	.53	.47	.50	.52	.39	.44	.40	.45	.44	.52	.48
D12	-	-	-	-	-	-	.46	.55	.33	.52	.43	.50	.41	.43	.37	.45	.37	.40	.43	.41
J13	-	-	-	-	-	-	-	.37	.52	.45	.55	.44	.51	.36	.42	.35	.54	.44	.44	.38
D13	-	-	-	-	-	-	-	-	.38	.57	.37	.49	.41	.38	.37	.40	.34	.46	.40	.43
J14	-	-	-	-	-	-	-	-	-	.44	.54	.42	.42	.44	.39	.36	.37	.33	.42	.41
D14	-	-	-	-	-	-	-	-	-	-	.46	.48	.49	.43	.46	.39	.47	.47	.47	.53
J15	-	-	-	-	-	-	-	-	-	-	-	.47	.46	.37	.42	.36	.49	.45	.30	.38
D15	-	-	-	-	-	-	-	-	-	-	-	-	.41	.43	.40	.41	.43	.48	.45	.42
J16	-	-	-	-	-	-	-	-	-	-	-	-	-	.34	.47	.33	.49	.39	.42	.42
D16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.35	.47	.34	.52	.35	.48
J17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.29	.36	.35	.45	.40
D17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.32	.47	.26	.43
J18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.33	.44	.36
D18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.38	.47
J19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	.40
D19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note: FTPR - First-time pass rate; PNLE-Philippine nurse licensure examination; J-June; D-December; Year is indicated by its last two digits (e.g., 10 = 2010); All  $r_s$  are significant at  $p < .01$ .

**Table 5.** Selected 95% Confidence Intervals (CI) and R<sup>2</sup> of PNLE Results from 2010 to 2019

PNLE	Correlation Coefficient (95% CI) p value	R <sup>2</sup>
2010 and 2011	.809 (95% CI = .774, .839) $p < .01$	65.4%
2014 and 2015	.644 (95% CI = .583, .698) $p < .01$	41.5%
2018 and 2019	.588 (95% CI = .512, .655) $p < .01$	34.6%
June 2010 and December 2010	.510 (95% CI = .433, .579) $p < .01$	26.0%
June 2011 and June 2011	.624 (95% CI = .560, .681) $p < .01$	28.9%
December 2012 and December 2013	.550 (95% CI = .478, .615) $p < .01$	30.3%
June 2014 and December 2019	.414 (95% CI = .310, .509) $p < .01$	17.1%

Note: R<sup>2</sup> = percentage of explained variance;  $p$  =  $p$ -value

## Discussion

### Trend and Level of PNLE Performance of Nursing Schools

The trend in PNLE FTPR has steadily increased from 2010 to 2019, which reflects an improving quality of nursing education in the country. This finding addresses previous studies about concerns about the quality of Philippine nursing education (Bautista et al., 2018; Montegrigo, 2020; Rosales et al., 2014). With almost half of the nursing schools in this study achieving average to high levels of performance based on their 10-year PNLE FTPR, this indicates a positive development in Philippine nursing education, because FTPR is a major criterion in evaluating nursing program quality (CHED, 2017; NCSBN, n.d). This improvement in FTPR can be attributed to quality improvement initiatives (i.e., curriculum revisions) (CHED, 2009, 2017, 2019).

The significant reduction in PNLE examinees from 2010 to 2016 can be attributed to the decrease in enrollment in nursing schools in the early 2000s, which was influenced by two major factors. First, the socioeconomic and political factors in the U.S. resulted in visa retrogression. Fewer demand for international nurses in the U.S. has a tremendous impact on enrollment in Philippine nursing schools (Arends-Kuenning et al., 2015). Secondly, the implementation of the K-12 program to strengthen basic education in the Philippines created a gap in tertiary education enrollment, thus reducing the number of nursing graduates and subsequent PNLE examinees (Bringula et al., 2018). Conversely, the slight increase in PNLE examinees from 2017 to 2019 is related to the increased demand for international nurses in the U.S. during the preceding years.

### Reliability of PNLE FTPR

A critical aspect of reliability is the ability of a measuring tool to consistently show scores over time. In educational research, a test demonstrates consistent results when administered repeatedly as a similar or parallel test in an index of reliability. Reliability is one requisite to validity, thus, licensure examinations must have an established reliability.

The findings in our study provide evidence of the high reliability of the PNLE, as evidenced by the statistically significant results across all examination schedules, whether taken semi-annually or annually. Though Foreman (2017) argues that high variability may suggest reliability issues, the moderate to strong correlation coefficients provide evidence of a high level of reliability of the PNLE. However, we suggest that the annual combined PNLE data should be used for high-stakes decision-making. The stronger evidence provided by combining the two PNLE as a single data eliminates potential measurement errors related to PNLE FTPR reporting. When official PNLE results are posted, schools with no examinees are reported as having zero FTPR, which is similar to schools with examinees but achieved zero FTPR. If all graduates of a particular school of nursing take and pass the May/June PNLE, there will be no examinees in the November/December PNLE, which may be reported as zero FTPR.

The wide variability and weaker correlation between semi-annual PNLE FTPR is likely related to the selection bias of nursing schools when determining which graduates will take the PNLE. The decision to allow nursing graduates to take the PNLE is commonly determined by, other than academic performance, satisfactory completion of comprehensive pre-



board examinations, and nursing review classes provided by either in-house faculty or outsourced review centers. As such, students who are stronger academically are more likely to take and pass the PNLE, leaving the academically weaker students to take the next PNLE exam. Thus, the PNLE FTPR of the two groups of pre-selected graduates may show wide variability. Moreover, nursing schools have preferences on whether to take the PNLE in May/June or November/December. Nursing schools may allocate more time to prepare their graduates for the PNLE that even when graduation takes place in April, they may spend another six months and then take the PNLE in November/December. This practice has helped schools ensure a high FTPR.

A methodological limitation that we acknowledge in this study is how missing data was treated and analyzed. As previously mentioned, schools with no examinees were left blank and no imputations were made. From the total 517 nursing schools, correlation coefficients were computed based on 262 to 487 schools, which were more than the required sample size of 82 based on power analysis. We initially considered mean substitution as an imputation alternative to resolve the missing data but there had been inconsistent perspectives on the use of this technique (Tabachnick & Fidell, 2013). Furthermore, although multiple imputation is currently the most sophisticated strategy to address missing data, in this case - MAR, this can also lead to bias in statistical analysis (van Ginkel et al., 2020). Since statistical power is dependent on the sample size (Polit, 2010; Tabachnick & Fidell, 2013), we consider our findings of sufficient statistical power despite the missing data.

Finally, considering that testing is a multifactorial phenomenon, other variables should not be discounted that may also be considered for wide variability in FTPR since 20 similar versions of the PNLE were administered in 10 years. Differences in the profile of nursing graduates, changes in curriculum, teaching, faculty composition, institutional leadership, PNLE test writers, test administration processes, scoring, and environmental factors may have potentially accounted for random and systematic measurement errors, causing wide variations in PNLE FTPR (McCoy, 2019; Waltz et al., 2017).

### Conclusion and Recommendations

This research is the first of its kind to determine the reliability of the PNLE. Based on the research findings, we conclude that the PNLE is a highly reliable tool to measure the competencies of nursing graduates and to determine the quality of nursing education programs. Schools of nursing should ensure consistent implementation of admission, progression, retention, graduation, and licensure preparation policies to attain and maintain effective structures and processes that produce

satisfactory outcomes, such as PNLE performance. With the wide variability and multidimensionality of factors influencing PNLE outcomes, regulatory agencies should be cautious in gauging the quality of a nursing program based on one fixed measurement tool. Nevertheless, the usefulness of the PNLE as a measurement tool to determine the quality of nursing programs may be enhanced by analyzing the annual combined PNLE results to assess the performance of the entire cohort of nursing graduates, compared to analyzing the performance of their graduates in two different PNLE exams.

Limitation in this study includes the number of MAR data of schools that had no examinees in certain PNLE. Despite the presence of MAR data, the sample size was sufficient to produce the desired power and effect size. Future studies may include imputation before data analysis. Since this is breakthrough research on PNLE reliability, the conduct of similar research is recommended to strengthen the evidence on the reliability of the PNLE.

### References

- Archer, J., Lynn, N., Coombes, L., Roberts, M., Gale, T., Price, T., & de Bere, S.R. (2016). The impact of large scale licensing examinations in highly developed countries: a systematic review. *BMC Medical Education*, 16(212); DOI: 10.1186/s12909-016-0729-7.
- Arends-Kuenning, M., Calara, A., & Go, S. (2015). International migration opportunities and occupational choice: a case study of Philippine nurses 2002 to 2014. Paper presented at the Institute for the Study of Labor. Bonn, Germany.
- Bautista, J., Ducanes, G., & David, C. (2018). Quality of nursing schools in the Philippines: Trends and evidence from the 2010-2016 Nurse Licensure Examination results. *Nursing Outlook*, 00, 1-11. doi.org/10.1016/j.outlook.2018.12.012
- Board of Nursing (BON, 2017). Syllabi of the Nurse Licensure Examination. Professional Regulation Commission, Republic of the Philippines. www.prc.gov.ph/nursing
- Bringula, R., Balcoba, A., Alfaro, L. & Merritt, J. (2018). Managing the perceived impact of K to 12 implementation on academic tenure and financial stability: evidence from five higher education institutions in the Philippines. *Educational Research for Policy and Practice*.doi.org/10.1007/s10671-018-9239-1
- Commission on Higher Education (2017). (Memo) Memorandum Order No. 15. Quezon City.
- CHED CMO 18, s. 2011. Amendments to Article XI - Sanctions of CMO 11., s. 2009. https://ched.gov.ph/wp-

- content/uploads/2017/10/CMO-No.18-s2011.pdf  
Federation of State Boards of Physical Therapy (FSBPT). (n.d.). Ensuring validity. <https://www.fsbpt.org/Free-Resources/NPTE-Development/Ensuring-Validity>
- Foreman, S. (2017). The accuracy of state NCLEX-RN passing standards for nursing programs. *Nurse Education Today*, 52, 81-86. DOI: <http://dx.doi.org/10.1016/j.nedt.2017.02.019>
- Foreman, S. (2019). Stability and validity of NCLEX-RN state pass rate standards. *Nursing Education Perspectives*, 40(1), E3-E8. DOI: 10.1097/01.NEP.0000000000000412
- Furr, R. M. & Bacharach, V.R. (2014). *Psychometrics: An introduction*, 2nd ed. Thousand Oaks, CA: Sage.
- Keller, S.P. & Kelvin, E.A. (2013). *Munro's Statistical Methods for Health Care Research*, 6th ed., Wolters-Kluwer Health, Lippincott Williams, and Wilkins. Philadelphia, PA.
- Masselink, L. & Lee, S. (2010). Nurses, Inc.: expansion and commercialization of nursing education in the Philippines. *Social Science & Medicine*, 71, 166-172.
- McCoy, T. (2019). *Statistical Methods II (Lecture)*. University of North Carolina at Greensboro. Greensboro, North Carolina.
- Montegrigo, J. (2020). A five-year trend analysis of the Philippine nurse licensure examination (2014-2018). *Philippine Journal of Nursing*, 89(2), 47-51.
- National Council State Boards of Nursing. (NCSBN, 2019). Reliability of the NCLEX examinations. <https://ncsbn.org/publications/reliability-and-validity>.
- NCSBN. (n.d.). Reliability of the NCLEX examinations. <https://www.ncsbn.org/4657.htm>
- National League for Nursing. (NLN, 2012). The fair testing imperative in nursing education. [www.nln.org/newsroom/nln-position-documents/nln-living-documents](http://www.nln.org/newsroom/nln-position-documents/nln-living-documents)
- Polit, D. (2010). *Statistics and data analysis for nursing research*, 2nd ed. Pearson Education Inc. New Jersey.
- PRC BON (2019). Program of the Nurse Licensure Examination. [https://www.prc.gov.ph/sites/default/files/NU\\_RSE\\_boardprogram\\_Nov2019.pdf](https://www.prc.gov.ph/sites/default/files/NU_RSE_boardprogram_Nov2019.pdf).
- Rosales, A., Arugay, Y., Divinagracia, C., & Castro-Palaganas, E. (2014). Analytical study of The nurses' licensure examination performance of graduates of Philippine colleges of nursing. *Philippine Journal of Nursing*, 84 (1), 4-23.
- Schober, P., Boer, C., & Schwoarte, L.A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia and Analgesia*, 125(5), 1763-1768. DOI:10.1213/ANE.0000000000002864
- Tabachnick, B.G. & Fidell, L.S. (2013). *Using multivariate statistics*, 6th ed. Pearson Education, Inc. New Jersey.
- van Ginkel, J.R., Linting, M., Rippe, R.C., & van der Voort, A. (2020) Rebutting existing misconceptions about multiple imputation as a method for handling missing data. *Journal of Personality Assessment*, 102(3), 297-308. DOI: 10.1080/00223891.2018.1530680

Waltz, C., Strickland, O., & Lenz, E. (2017). *Measurement in nursing and health research*. New York: Springer Publishing Company, LLC.

## ABOUT THE AUTHORS



**James Montegrigo**, PhD, RN, CMSRN, is an Assistant Professor, Honors Faculty, and the Academic Success Program Coordinator of the University of North Carolina at Charlotte School of Nursing. He earned his Doctor of Philosophy in Nursing Science from the University of North Carolina at Greensboro. He completed his Bachelor of Science in Nursing at the University of Santo Tomas as a Magna Cum Laude graduate and academic scholar, and Master of Science in Nursing (major in Adult Health Nursing) at Saint Paul University Philippines, Tuguegarao City.

Dr. Montegrigo's research interests include nursing education, nursing workforce, and HIV/AIDS. He has published his research at national and international journals, and presented his scholarly works at various regional, national, and international conferences. He is a peer reviewer of various international nursing journals such as the *Journal of Neuroscience Nursing*, *Nursing Education Perspectives*, *Journal of Community Health Nursing*, and *Journal of Professional Nursing*. He is currently a member of the Test Development Committee of the Medical-Surgical Nurse Certification Board of the AMSN. He is a faculty counselor of Sigma Theta Tau International Honor Society of Nursing - Gamma Iota Chapter and a member of the Honor Society for International Scholars (Phi Beta Delta Chapter).



**Geraldine Y. Ferreras**, MSN, RN is an Assistant Professor in School of Health and Natural Sciences at Saint Mary's University, Bayombong, Nueva Vizcaya. She has been an undergraduate faculty member since January 2006 and also in the Graduate School since June 2008. Her research is focused on developing competent and compassionate nurses.



**Kelly Powers**, PhD, RN, CNE, FNAP is an Associate Professor in the School of Nursing at the University of North Carolina at Charlotte. She has been a faculty member at UNC Charlotte since 2008 and has taught in the BSN, MSN, and DNP programs. Dr. Powers' research is focused on developing and testing innovative interventions, such as simulation and interprofessional education, to improve the practice of nurses and other healthcare professionals.