Technical Manual

Nursing Entrance Exam (NEX)

National League for Nursing
2600 Virginia Avenue NW, 8th Floor
Washington, DC 20037

All rights reserved. No part of this manual may be reproduced in print or by photocopy, or in any other manner, without the express written permission of the National League for Nursing.

March 2024

Table of Contents

List of Tables 4
List of Figures 5
Preface 6
1. The National League for Nursing 7
1.1 NLN Mission 7
1.2 Overview of the NLN 7
2. NEX Overview 8
2.1 Purpose and Intended Population 8
2.2 Objectives 8
2.3 Test Overview 9
2.4 Changes from PAX to NEX 10
3. Test Development and Item Writing 13
3.1 Selection and Training of Item Writers 14
3.2 Item Review Before Administration 14
3.3 Pretesting Items 14
3.4 Review of Items After Administration 15
3.5 Delivery of Test Forms 15
4. Test Specifications 15
4.1 Statistical Specifications 16
4.1.1 Statistical Terms 16
4.1.2 Test and Item Specifications 16
4.2 Content Specification Updates 17
4.3 Verbal Ability Test: Objectives, Blueprint, and Statistical Specifications 17
4.3.1 Word Knowledge 17
4.3.2 Reading Comprehension 18
4.4 Mathematics Test: Objectives, Blueprint, and Statistical Specifications 19
4.5 Science Test: Objectives, Blueprint, and Statistical Specifications 22
5. Psychometric Quality 24
5.1 Validity 25
5.1.1. Proposition 1: Academic Abilities are Pre-requisites to Success in a Nursing Program 25

5.1.2 Proposition 2: Test Content Aligned to Pre-requisite Abilities 26
5.1.3 Proposition 3: Defining Success Metrics in Nursing Programs 27
5.1.4 Proposition 4: Evidence of Relationship with Student Performance 27
5.2 Reliability 28
6. Scales, Norms, and Linking 28
6.1 Scaling and Percentile Ranks 28
6.2 Norms 29
6.2.1 Methodology for Norming 30
6.2.1 Future Operational Norms 31
6.3 Linking 31
6.3.1 Linking Methodology 31
6.3.2 Linking Considerations 32
7. Reporting and Quality Controls 32
7.1 Participant Report 33
7.2 Institutional Coaching Report 34
7.3 Institutional Report 34
7.4 Interpreting and Using Scores 35
7.5 Quality Control for Continuous Mode Tests 35
7.5.1 Quality Control During Test Administration 36
7.5.2 Quality Control in Test and Item Analysis 37
8. Fairness and Equity in Testing 38
9. Other Information Regarding NLN Test Administration 38
9.1 Limitations of Exams 38
9.2 Retake Policy 39
9.3 Challenge Items 39
9.4 Test Score Retention 39
9.5 Fair Testing and Ethics Statement 39
References 41
Appendix A 43
Appendix B 44
Appendix C 45

Appendix D 52
Appendix E 55

List of Tables

Table 1. High Level NEX Blueprint 9
Table 2. Summary of Key Changes to the PAX Exam 10
Table 3. NEX Word Knowledge Blueprint by Difficulty Level 18
Table 4. NEX Reading Comprehension Blueprint Distribution by Topic and Cognitive Dimension 18
Table 5. NEX Verbal Ability Content Outline 19
Table 6. NEX Mathematics Blueprint 21
Table 7. NEX Mathematics Content Outline 21
Table 8. NEX Science Blueprint 23
Table 9. NEX Science Content Outline 24
Table 10. Examples of Prior Predictive Validity Studies 27
Table 11. NEX Test Statistics 28
Table 12. NLN Quality Controls During Test Administration 36
Table 13. NLN Quality Controls During Test Analysis 37

List of Figures

Figure 1. NEX Participant Report Sample 33
Figure 2. NEX Institutional Coaching Report Sample 34
Figure 3. Institutional Report Sample 35

Preface

This manual contains technical information about the National League for Nursing (NLN) Nursing Entrance Exam (NEX). The purpose of this technical manual is to provide content and technical characteristics of the exam, test development information, and information gathered to date for the valid and appropriate interpretations and use of test scores. By providing this information the user can determine if this exam meets their needs and how they will in turn choose to use the results. Additionally, the information can help institutions transition from using the Pre-Admission Exam (NLN Testing Services, 2016) to the NEX.

The content of this manual intends to follow the guidelines set forth by the testing industry as stated by the Standards for Educational and Psychological Testing (APA, AERA, & NCME, 2014), and the Fair Testing Guidelines for Nursing Education (NLN, 2020.)

1. The National League for Nursing

1.1 NLN Mission

The National League for Nursing (NLN) promotes excellence in nursing education to build a strong and diverse nursing workforce to advance the health of our nation and the global community.

1.2 Overview of the NLN

Dedicated to excellence in nursing, the NLN is the premier organization for nurse faculty and leaders in nursing education. The NLN offers professional development, networking opportunities, testing services, nursing research grants, and public policy initiatives to its 40,000 individual and 1,200 institutional members. NLN members represent nursing education programs across the spectrum of higher education and health care organizations and agencies.

Founded in 1893 as the American Society of Superintendents of Training Schools for Nurses, the NLN was the first nursing organization in the United States. Headquartered in Washington, D.C., the NLN is led by a board of governors elected by membership for three-year terms. The volunteer board chair works with the NLN's president and chief executive officer.

2. NEX Overview

2.1 Purpose and Intended Population

The primary purpose of the NEX is to provide educators with a standardized instrument for evaluating the academic ability of applicants to nursing programs. Results from the exam can be used as one component in the total evaluation of each individual applicant in admission and student support decisions.

The NEX is intended for students seeking admission to a pre-licensure nursing program. Nursing programs vary in their admission practices and requirements, including when students are required to complete the exam (e.g., before or after completing required courses) and submit their scores. As such, students taking the exam may span a range of educational preparedness. Admissions administrators and faculty within a nursing program are responsible for determining the appropriateness of this exam for their program. The content and statistical specifications included in this manual can be useful in making the decision.

2.2 Objectives

General objectives of the exam are to:

1) Measure applicants' fundamental academic competencies in verbal, mathematics, and science abilities;
2) Provide nursing programs valid, reliable data as one measure to guide admission decisions; and
3) Provide nursing programs valid, reliable data as one measure to inform student support and success efforts.

The Verbal Ability test is designed to measure two primary skills: word knowledge and reading comprehension encountered at the high-school or early college level. The Mathematics test is designed to measure an applicant's ability to solve basic computation and mathematics problems no higher than Algebra I (approximately the tenth-grade high-school level) with a focus on skills encountered in the nursing curriculum, including numbers and operations, measurement, algebra, and data & information. The Science test is designed to measure an applicant's knowledge of general principles of high-school or early college level chemistry, biology, human anatomy, human physiology, and health. More specific objectives for each component of the exam are provided in the respective sections later in this document.

2.3 Test Overview

The NEX is a battery of three computer-based tests delivered in a single administration with a live proctor, either in-person (on-site at an institution's approved testing center) or remotely (with a contracted live proctor monitoring the administration virtually). All test items are multiple-choice format with four response options. Items are scored as correct or incorrect without penalty for guessing. Unanswered items are counted as incorrect. All tests must be completed to generate a score report.

Each test consists of both operational (scored) and pretest (unscored) items that have undergone content and sensitivity reviews; however, only the operational items are scored and used to assess student achievement. Students are allowed 60 minutes to complete each test broken down as follows:

- Verbal Ability: 58 items (50 operational and 8 pretest)
    - 25 word knowledge and 25 reading comprehension
- Mathematics: 45 items (40 operational and 5 pretest)
- Science: 60 items (55 operational and 5 pretest)

Table 1. High Level NEX Blueprint

		Number of Scored Items	Unscored Items	Percent of Total Test
Verbal		**50**	**8**	**34%**
	Word Knowledge	25		17%
	Reading Comprehension	25		17%
Math		**40**	**5**	**28%**
	Numbers and Operations	12		8%
	Measurement (Conversions)	14		10%
	Algebra	7		5%
	Data & Information	7		5%
Science		**55**	**5**	**38%**
	Biology	20		14%
	Human Anatomy	11		8%
	Human Physiology	11		8%
	Chemistry	5		3%
	Health	8		6%
Total		**145**	**18**	**100%**

Note. Numbers may not equal 100% due to rounding.

2.4 Changes from PAX to NEX

As part of the NLN Assessment Service Division's commitment to continuous improvement, two consultants, a psychometric evaluator and expert nurse educator, conducted an extensive review of the 2016 Pre-Admission Exam (PAX) (NLN Testing Services) including a review of the technical manual, blueprint, operational and trend data, implementation practices, and score reports. Based on their findings and considerable inputs from the NLN's lead nursing committee, communities of interest, and Subject Matter Experts (SMEs), representing diverse institutions and nursing programs, the NLN decided to update and rebrand the PAX to the Nursing Entrance Exam (NEX). The summary of key updates, depicted in Table 2, is intended to provide nursing program faculties and administrators currently using the PAX with relevant information about the substantive changes made to the exam as they plan for transitions to the NEX[1]. Appendix A contains a list of the contributors.

Table 2. Summary of Key Changes to the PAX Exam

Change	PAX	NEX	Improvements to Validity
1. Test Delivery and Administration			
1.1 Time Allotment	40 minutes per section	60 minutes per section	Reduce effects of "speededness"
1.2 Test Type	Random Parallel Test	Modified Linear-On-The Fly Test	Blueprint and Item Specifications drive test delivery
1.3 Virtual Proctoring	2016 – Onsite only 2020 – Onsite and Virtual using "record-and-review" proctor	Onsite or Virtual with provisions to use live virtual proctor only	Test Security and comparability of scores across administration methods
2. Blueprint/Content Specifications			
2.1 Number of Items	Verbal: 60 Math: 40 Science: 60	Verbal: 50 plus 8 pre-test Math: 40 plus 5 pre-test Science: 55 plus 5 pre-test	Pre-test items (unscored) help deepen item bank to increase test security
2.2 Verbal Ability Test	Reading passages replaced in 2021 **Reading Comprehension**: 30 items, 6 reading passages	Re-norm test reflecting 2021 passages **Reading Comprehension**: 25 items, 5 passages (assess	Implement feedback to reflect academic Abilities needed for success in a nursing

[1] Standard 7.14 - When substantial changes are made to a test, the test's documentation should be amended, supplemented, or revised to keep information for users current and to provide useful additional information or cautions

Change	PAX	NEX	Improvements to Validity
	(assess comprehension, application, inference) **Word Knowledge:** 30 items (assess vocabulary)	comprehension, application, inference) **Word Knowledge:** 25 items (assess vocabulary and deductive reasoning through context clues and prefixes/suffixes/roots)	program.
2.3 Science Test	**Physics**: 5 items **Biology**: 20 items **Chemistry**: 13 items **Anatomy & Physiology**: 14 items **Health**: 8 items	**Physics**: 0 items **Biology**: 20 items **Chemistry**: 5 items **Anatomy:** 11 items **Physiology**: 11 items **Health**: 8 items	Implement feedback to reflect academic abilities needed for success in a nursing program.
2.4 Mathematics Test	**Numbers & Operations**: 29 items **Conversions**: 9 items **Algebra**: 5 items **Mathematical Concepts**: 4 items **Geometry**: 3 items	**Numbers & Operations**: 12 items **Measurement**: 14 items **Algebra**: 7 items **Data & Information**: 7 items	Implement feedback to reflect academic abilities needed for success in a nursing program.
3. Statistical Specifications			
3.1 Test and Items	Target Average Item Difficulty Level: .60 Target Average Discrimination Index (point biserial): .40	Target Average Item Difficulty Level: .63 Item Difficulty Range: .20 - .89, Target Average Discrimination Index (point biserial): .40, Minimum: .20	Improve item statistics and test reliability
4. Scale and Reporting			
4.1 Test Scale	Composite score scaled as 0 – 200, computed as summation of raw scores from each section with Mathematics weighted at 1.5.	Composite score scaled as 0-300, computed as summation of percentile ranks from each section, weighted equally	Implement feedback of equal weighting and desire for percentile ranks for ease of score interpretation.
4.2 Student Reports	Composite score, composite percentile rank, and raw	Composite score (i.e., composite percentile),	Meaningful interpretation and use

Change	PAX	NEX	Improvements to Validity
	scores for sections	subject test raw scores (reported as percent correct), subject test percentile ranks, interpretive verbiage about the meaning of scores	of scores (student's performance relative to norming population, student's relative strengths and weaknesses – e.g., math skills stronger than verbal skills based on percentile ranks)
4.3 Coaching Reports	Raw (percent scores) at two levels organized across content labeled RN or PN with limited meaningfulness to student or nursing program	Raw (percent scores) for each section test (verbal, math, science) and each sub-content area (e.g., reading comprehension, word knowledge, biology, chemistry, algebra, etc.)	Meaningful interpretation and use of scores for student remediation
4.4 Institutional Reports	Composite percentile and raw scores for section tests	Composite exam score and composite and subject test percentiles	Meaningful interpretation and use of scores; cohort planning or remediation (e.g., respond to cohort math skills)
5. **Preparatory Materials for Remediation**			
5.1 Prep Materials	Online practice exam and content review publication	Online practice exam, diagnostic assessment report, and review publication	Meaningful interpretation and use of scores for student remediation in each sub-content area
5.2 Practice exam delivery	Random selection of items from one subject item pool (e.g., Science)	Random selection of items from sub-content areas (e.g., biology)	Practice test follows blueprint

3. Test Development and Item Writing

The NLN follows the guidelines outlined in *Standards for Educational and Psychological Testing* (American Psychological Association [APA], American Educational Research Association [AERA], , & National Council for Measurement in Education [NCME], 2014) in developing, evaluating, revising, and maintaining its exams. The NLN Test Development Process (Appendix B) fosters adherence to the guidelines outlined in *the Standards* and the Fair Testing Guidelines for Nursing Education (National League for Nursing, 2012), including content reviews, item development processes, fairness and sensitivity reviews, and psychometric reviews. All steps in the test development process are intended to promote valid score interpretation and minimize the impact of construct irrelevance[2]. A summary of the test development process is delineated below as steps; however, the process of test development is iterative.

1. Target population and test purpose is defined.
2. Constructs to be assessed are clearly articulated and reviewed at the Subject Matter Expert (SME) and committee level.
3. Blueprints are developed and reviewed at the SME and committee levels. Content details and existing test items are reviewed for alignment.
4. Item writers are trained on the target population, blueprint, and item quality and submit items according to assignments defined by the blueprints.
5. Submitted items undergo an extensive review process, including content, sensitivity and bias, and editorial reviews to ensure fairness and non-bias.
6. Items are pretested and undergo a psychometric review for item quality and fairness.
7. Reliability estimates are computed for each section test and disaggregated as necessary for identified subgroups.
8. User norms are collected and analyzed for representation and potential reference groups.
9. Test delivery is standardized, includes clear and simple instructions and procedures, and is flexible to accommodations.
10. Summary data on examinee performance is provided to enable test users to find an examinee's relative standing within clearly defined reference groups.
11. Research is conducted on the predictive effectiveness of the tests and on a wide variety of topics to continue to build evidence of validity for the appropriate use and interpretation of the test scores.

[2] Standard 3.0 – All steps in the testing process, including test design, validation, development, administration, and scoring procedures, should be designed in such a manner as to minimize construct-irrelevant variance and to promote valid score interpretations for the intended uses for all examinees in the intended population.

3.1 Selection and Training of Item Writers

Items for the exam are formulated by item writers who are specialists in the types of skills measured according to the test blueprint. These item writers typically have graduate degrees in relevant areas and may have taught at the high school, college, or university level. NLN recognizes the diversity of applicants to nursing schools and is sensitive to selecting, insofar as is possible, item writers and reviewers from minority groups. Item writers receive training and resources enabling them to create quality items that are educationally relevant and meet the blueprint specifications. The collective review process further ensures item quality, consistency, fairness, and achieving the desired psychometric benchmarks.

3.2 Item Review Before Administration

Assessments should provide all test takers an equal opportunity to be able to show what they know and can do. This opportunity requires deliberate reviews of the content, wording, and statistical characteristics of each item. Items judged acceptable by all reviewers become eligible for pretesting. Item reviews include:

- Content: focus on accuracy, blueprint alignment, sub-content alignment, and correct answer rationale
- Bias and sensitivity: focus on item review through the lens of potential race, ethnicity, culture, gender, region, religion, and socioeconomic status and avoiding sensitive topics such as abortion, abuse, climate change, contraception, deportation of immigrants, ethnic disputes, euthanasia, killing of animals for sport, occult, political or religious disputes, prayer in school, rape, sexual behavior or innuendos, suicide, or torture.
- Editorial: focus on clear, concise, and consistent format and style, and correct grammar and language usage.

3.3 Pretesting Items

After content, sensitivity, and editorial reviews of items are completed, the items are used in an operational test form as pretest items, which are unscored items seamlessly integrated in the test to gather information regarding item performance. After an item has been pretested, a preliminary item analysis is conducted to provide information about the difficulty of the item, the relative attractiveness of the answer choices, and how well the item discriminated among examinees with various levels of ability. Another analysis, differential item functioning (DIF), provides information about the extent to which an item functions in the same way for different subgroups of examinees who have been matched on ability. Items that are deemed unacceptable may include items that are too difficult ($\geq$ 0.90) or too easy (<0.20), items that fail to distinguish examinees of high and low developed ability as measured by

the exam scores, and items on which there are large differences in performance for subgroups of examinees. These unacceptable items are eliminated from the item pool or revised and pretested again. Items that do not perform in an acceptable manner may be reviewed by test developers later to determine the possible cause(s) of poor performance.

3.4 Review of Items After Administration

The NLN's quality control process (Section 7.6) outlines a systemic review of the exam to ensure it performs as intended. Even with these processes in place, a faculty member or an examinee may occasionally raise a question about a test item at or after the test administration. These queries are conveyed to test developers, who carefully review the relevant item(s) and prepare a response to the inquirer that addresses the issue(s) noted. On rare occasions, an item may be rescored as the result of such a challenge.

3.5 Delivery of Test Forms

Multiple test forms are needed to maintain the security of a high-volume standardized continuous mode test (CMT) and subsequently sustain the validity of test scores. This exam employs a modified Linear-On-The-Fly-Test (LOFT) methodology to achieve that need. LOFT delivers unique forms that meet the blueprint specifications and pre-defined item and test parameters (Becker & Bergstrom, 2013). In the NEX, the test delivery system meets the blueprint specification and a set number of items in each pre-defined item difficulty category. Items are categorized by their difficulty level as follows: low, p = .70 - .89; medium, p = .50 -.69, and high, p = .20 - .49. The test delivery platform is programmed to randomly select items at three intersection points of the blueprint – sub-content, dimension, and item difficulty level. For example, in the Science test, the blueprint specifies three Biology items (sub-content) for *interpretation* (dimension) and one of those items will be of high difficulty. If four items meet all three of those requirements, the test delivery system will randomly select one of the four items available. This random selection process is repeated for each topic/dimension/item difficulty intersection, and in so doing, multiple comparable test forms are created "on-the-fly". This approach ensures all students receive the same number of items from each sub-content area, dimension, and difficulty level. For a more detailed discussion of simulations, comparability, and score interpretation under a modified LOFT methodology, reference Section 6.

4. Test Specifications

The exam specifications set forth the requirements for content and statistical characteristics that apply to every NEX test. The original Pre-Admission Exam (PAX) blueprints laid out the content

specifications based on a review of nursing textbooks and curricula, as well as a survey of educators actively working with nursing students, high school students, and the general college student population, to provide the basis for determining the prerequisite knowledge and appropriate level of verbal, mathematical, and science ability needed for students to succeed in a nursing program (NLN Pre-Admission Examination-RN™ Technical Manual, 2002). Building on that excellent work, evaluators, committees, and SMEs, conducted a thorough review of the PAX blueprint and recommended changes to update the exam to represent current high school, early college, and nursing curricula.

4.1 Statistical Specifications

A range of item difficulties is required to differentiate among examinees at various levels of abilities across subjects. Over time, changes in test delivery, teaching and learning, item exposure, or current events (e.g., Covid-19) can shift item difficulty or discrimination, and subsequently, the norming comparisons established for score interpretation and use. Monitoring and addressing statistical specifications contribute to maintaining the psychometric quality of the exam (Section 7.6).

4.1.1 Statistical Terms

Item Difficulty is the relative frequency with which examinees choose the correct response (Thorndike et al., 1991). Item difficulty ranges between 0.0 and 1.0 and can be described as the percentage of examinees who answer an item correctly, and is noted as *p.*

Item Discrimination, or *Point Biserial,* is used to differentiate performance between high- and low-performing examinees. Item discrimination ranges between -1.0 and 1.0 and measures the strength and direction of the total score of examinees with their performance (correct or incorrect) on a particular item (Yoo & Hambelton, 2019), and is noted as r_{pbs}.

4.1.2 Test and Item Specifications

The following statistical specifications are set forth: (a) the target average difficulty level of each subject test is $p = 0.63$, appropriate for a norm-referenced test comprised of multiple choice items with four response options (Lord, 1952; Matlock-Hetzel, 1997); (b) the acceptable item difficulty level ranges of $p = 0.20$ to $p = 0.89$, inclusive of the most desirable item difficulties levels ($p = 0.30$ - 0.70) for norm-referenced tests (Oerman & Gaberson, 2021); and (c) a minimum item discrimination threshold of $r_{pbs} = 0.20$ (Ebel & Frisbie, 1986; Oermann & Gaberson, 2021). By establishing these specifications, the NLN assembles each section test and its associated delivery to optimize reliability and validity of the use and interpretation of test scores. Items from the PAX item bank that meet those specifications populate the first set of NEX tests.

4.2 Content Specification Updates

Content and blueprint reviews from expert nursing educators, nursing advisory committees representing diverse institutions, and general education subject matter experts (SMEs) informed the NEX blueprint. Using the PAX blueprint as a reference point, the following changes are set forth to improve content validity: (a) removal of physics content from the Science and Mathematics sections as it not a pre-requisite in nursing programs nor a consistent high school requirement for graduation; (b) removal of extended geometry items (e.g., circumference, radius, diameter, angles), content not typically evident in nursing curricula; (c) limit remaining geometry problems to assess application of given formulas, measurement, and scale; (d) increase representation of conversion problems to align with the nursing curriculum for measurement conversions while decreasing representation of computational problems in numbers and operations as all students have access to a built-in calculator while testing; and (e) add data & information to assess basic descriptive statistics and interpretation of information presented in tables or graphs.

4.3 Verbal Ability Test: Objectives, Blueprint, and Statistical Specifications

The Verbal Ability test consists of two sections: word knowledge (25 items) and reading comprehension (25 items) representing skill levels often expected in high school or the first year of college. The word knowledge section assesses the examinee's ability to deduce the meaning of a word through context clues provided in a sentence or use of prefix/suffix/root knowledge and apply existing vocabulary to identify synonyms or antonyms. The reading comprehension subsection is composed of five informational passages with content of a scientific, health, technology, or general nature suitable for measuring the examinee's ability to understand written materials at a ninth-grade level or higher. Reading comprehension test items explore three dimensions across the topical passages; comprehension (identifying details), application (including analysis, vocabulary, and main ideas), and making an inference.

Objectives of the Verbal Ability exam are to:

- Assess examinees' ability to identify the meaning of a word by applying existing word knowledge and deductive reasoning skills;
- Assess examinees' ability to comprehend, apply, and analyze informational passages from the disciplines of social, health, and natural sciences and technology by using comprehension skills, close reading strategies, deductive reasoning, and making inferences.

4.3.1 Word Knowledge

The first roll out of the NEX is built on the highest performing items from the PAX word knowledge item bank, evaluated using operational data from 2022. Word knowledge items have an

average difficulty level of p = 0.63 with 96% of the items within the desired difficulty range of p = 0.20 to p = 0.89; an average discrimination index, reported as the corrected item-total correlation, of r_{pbs} = 0.29; a range of r_{pbs} = 0.11 to r_{pbs} = 0.47 with 93% of the items meeting the desired minimum r_{pbs} = 0.20. Table 3 presents the NEX Word Knowledge blueprint and delineates the percentage of items at each difficulty level.

Table 3. NEX Word Knowledge Blueprint by Difficulty Level

Content Area: Word Knowledge				
Difficulty Level[1]	**Low**	**Medium**	**High**	**Total**
Total Items	4	15	6	25
Total %	16%	60%	24%	

[1]Item difficulty serves as a proxy for Grade Level 9-12+, the higher the difficulty, the likelihood of more advanced vocabulary.

4.3.2 Reading Comprehension

The first roll out of the NEX is built on the highest performing items from the PAX reading comprehension item bank, evaluated using operational data from 2022. A total of 5 passages and 25 items that most closely met the item difficulty and discrimination parameters were selected. Reading comprehension items have an average difficulty level of p = 0.68 with 88% of the items within the difficulty range of p = 0.20 to p = 0.89; an average discrimination index, reported as the corrected item-total correlation, of r_{pbs} = 0.29; with 92% of the items meeting the minimum r_{pbs} = 0.20.
Table 4 presents the NEX Reading Comprehension blueprint.

Table 4. NEX Reading Comprehension Blueprint Distribution by Topic and Cognitive Dimension

Content Area: Reading Comprehension				
Dimensions	**Comprehension**	**Application**	**Inference**	**Total**
Total Items	5	15	5	25
Total %	20%	60%	20%	

The Verbal Ability Content Outline, Table 5 below, provides additional detail of the general content areas to be sampled. A wide range of dimensions provides item writers and test developers with flexibility to sample content domain.

Table 5. NEX Verbal Ability Content Outline

Dimensions	Rationale for Testing Skill	Expected Outcome
Content Area: Word Knowledge (25)		
Context Clues	Vocabulary is viewed as a comprehension skill, where words are embedded within contextual clues (sentences) and word meanings are identified through use of deductive reasoning instead of rote memorization.	Examinees will primarily use their reading comprehension skills to help identify word meanings and may use other strategies as appropriate.
Suffix/Prefix/Root (Affixes)	Many words in the English language contain Greek and Latin affixes and most state ELA standards call for study of Greek and Latin roots, suffixes, and prefixes.	Examinees will primarily use their knowledge of affixes to help identify word meanings and may use other strategies as appropriate.
Synonyms	Strong word knowledge enables readers to identify words with similar meanings.	Examinees will use existing word knowledge to identify synonyms presented in the context of a sentence.
Antonyms	Strong word knowledge enables readers to identify words with opposite meanings.	Examinees will use existing word knowledge to identify antonyms.
Words Frequently Confused	A key vocabulary knowledge skill is distinguishing between meanings of words that sound similar but have different meanings.	Examinees will use their existing vocabulary to identify word meanings.
Content Area: Reading Comprehension (25)		
Comprehension	Reading comprehension entails identifying key details and main ideas within a given reading selection.	Examinees will show comprehension of reading passages by correctly identifying pertinent details from a given passage.
Application	Application requires readers to apply information from a given passage to similar situations and use deductive strategies to identify key word meanings.	Examinees will apply comprehension of reading passages to determine meanings of unfamiliar words or apply information from reading passages to new or different situations
Inference	Strong analysis and application of information enables readers to draw accurate inferences about information or details left unsaid in a given reading passage.	Examinees will use knowledge of information from reading passages to infer key information related but unstated in a given passage.

4.4 Mathematics Test: Objectives, Blueprint, and Statistical Specifications

The Mathematics test is designed to measure an applicant's competency in solving

basic arithmetic problems in (a) numbers & operations; (b) measurement; (c) algebra, and (d) data & information. This section includes computation and word problems which give applicants the opportunity to demonstrate their ability to perform basic mathematical operations as well as problem-solving and logic skills, reflective of the foundation needed for success in the nursing curriculum. Assessed content reaches as far back as Grades 6-8 and extends to 10th grade Algebra I. The importance of the content assessed is stated by the Common Core State Standards for Mathematics (National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010):

> …some of the highest priority content for college and career readiness comes from Grades 6-8. This body of material that includes powerfully useful proficiencies such as applying ratio reasoning in real-world and mathematical problems, computing fluently with positive and negative fractions and decimals, and solving real-world and mathematical problems involving angle measure, area, surface area and volume.

Objectives of the Mathematics exam are to assess applicants' abilities to:

1. Perform the four basic operations (addition, subtraction, multiplication, and division) on integers, decimals, fractions, and percentages in computation and problem-solving;
2. Convert one form of measurement to another while maintaining proportion or ratio and apply basic formulas to express measurements in computation and problem-solving;
3. Perform basic algebraic manipulations in computation and problem-solving;
4. Compute values to describe data and interpret summary of graphical information in computation and problem-solving.

The first roll out of the NEX is built on the highest performing items from the PAX mathematics item bank evaluated using operational data from 2022, and that met the content blueprint. Mathematics items have an average difficulty level of $p = 0.65$ with 95% of the items within the difficulty range of $p = 0.20$ to $p = 0.89$; an average discrimination index, reported as the corrected item-total correlation, of $r_{pbs} = 0.36$, a range of $r_{pbs} = 0.20$ to $r_{pbs} = 0.52$; and 100% of the items meet the minimum $r_{pbs} = .20$. Table 6 presents the Mathematics blueprint.

Table 6. NEX Mathematics Blueprint

	Dimensions		Content Area Totals	
	Computation	**Problem Solving**	**Items**	**Percentage**
Content Areas				
Numbers and Operations	7	5	12	30.0%
Measurement	3	11	14	35.0%
Algebra	3	4	7	17.5%
Data & Information	2	5	7	17.5%
Dimension Total Items	15	25	40	100.0%
Dimension Total %	38%	63%		

Table 7 provides a further delineation of the general content areas to be sampled when Mathematics section forms are assembled. A wide range of subtopics provides item writers and test developers with flexibility to sample content domain.

Table 7. NEX Mathematics Content Outline

Content Area	Rationale for Testing Skill	Expected Outcome by Dimension
Numbers and Operations	Rational and irrational numbers, their properties, and units comprise the basis for computation and quantification of real-life problems. Nursing students should demonstrate the ability to compute fluently with integers, decimals, and fractions (including percentages), figuring out when different forms are equal to each other, and solve real-world problems.	**Computation (5)**: Use rules and properties of real numbers to compute single and multi-step problems. **Logic and Problem Solving (7)**: Use rules and properties of real numbers and units to reason quantitatively to solve single and multi-step problems.
Measurement	Units and scale, introduced in grades 4 and ratios and proportional relationships, introduced in Grade 6, both practiced through high school, comprise the basis for accurate and meaningful measurements, relationships, and computations of real-life problems. Nursing students should demonstrate the ability to accurately measure, use formulas, compute, and convert within and between a variety of units and scales and in order to solve real-life problems.	**Computation (6)**: Use units, scales, ratios, and proportional relationships to compute solutions to measurement problems. **Logic and Problem Solving (8)**: Use units, scales, ratios, and proportional relationships to solve measurement problems in real-world contexts.
Algebra	The manipulation of expressions and equations, introduced in grade 6 and practiced through high school, form the basis for reasoning about solving real-life problems.	**Computation (3)**: Use algebraic concepts and principles to manipulate expressions, equations, and inequalities to compute solutions.

Content Area	Rationale for Testing Skill	Expected Outcome by Dimension
	Nursing students should demonstrate a facility with manipulating expressions and equations to compute and solve problems. The use of equations helps nursing students understand real-world problems with which they and their patients are faced. Proficiency with equations will aid in not only solving problems but also in helping their patients understand the results of their care plans.	**Logic and Problem Solving (4)**: Use algebraic concepts and principles in expressions, equations, and inequalities, to solve problems in real-world contexts.
Data and Information	Summary statistics, introduced in grade 6 and practiced through high school, describe shape, center, and spread of the data and are used to make informed decisions for real-life problems. Data plots present information by organizing it into a meaningful representation to help examine patterns, relationships, distributions, differences, and trends within or between data sets for informed decision-making. Nursing students should demonstrate the ability to compute and interpret basic summary statistics and interpret data and information recorded in a variety of tabular or graphical forms.	**Computation (3)**: Compute and compare the mean, median and mode(s) of real-value data sets. **Logic and Problem Solving (4)**: Use summary statistics of real-value data to compare two or more data sets. Interpret data plots or tabular information to identify meaning, patterns, relationships, or trends.

4.5 Science Test: Objectives, Blueprint, and Statistical Specifications

The Science test is designed to measure an applicant's competency in general science knowledge and principles that provide the foundation needed for success in the nursing curriculum. The content of this section is intended to be set primarily at the high-school level with some college 100-200 level content possible for the more difficult items. The subjects include (a) biology, (b) human anatomy, (c) human physiology, (d) chemistry, and (e) health. Specific dimensions of science include the cognitive skills of (a) knowledge/comprehension; (b) application of principles, basic research, and measurement; and (c) interpretation of charts, graphs, and diagrams.

Objectives of the Science exam are to assess participants' knowledge across biology, anatomy and physiology, chemistry, and health as follows:

1. Identify facts and terms (Knowledge);
2. Recognize processes and principles (Comprehension);
3. Apply principles to specific situations (Application);

4. Apply basic research and measurement (Application); and
5. Interpret charts, graphs, and diagrams (Interpretation)

The first roll out of the NEX is built on the highest performing items from the PAX science item bank evaluated using operational data from 2022, and that met the content blueprint. Science items have an average difficulty level of p = 0.61 with 99% of the items within the difficulty range of p = 0.20 to p = 0.89; an average discrimination index, reported as the corrected item-total correlation, of r_{pbs} = 0.26, a range of r_{pbs} = 0.09 to r_{pbs} = 0.45; and 88% of the items meet the minimum r_{pbs} =.20. Table 8 presents the Science blueprint.

Table 8. NEX Science Blueprint

	Dimensions			**Content Area Totals**	
	Knowledge/ Comprehension	**Application**	**Interpretation**	**Items**	**Percentage**
Content Areas:					
Biology	11	4	5	20	36%
Human Anatomy	8	1	2	11	20%
Human Physiology	2	5	4	11	20%
Chemistry	3	1	1	5	9%
Health	2	5	1	8	15%
Dimension Total Items	26	16	13	55	100%
Dimension Total %	47%	29%	24%		

Table 9 provides additional detail of the general content areas to be sampled from when the Science section forms are constructed. A wide range of topics provides item writers and test developers with flexibility to sample content domain.

Table 9. NEX Science Content Outline

Biology (20)
Cell structure and function (including diffusion and osmosis, mitosis and meiosis) Genetics and DNA Plants and photosynthesis Evolution - evidence, theories, and classification of organisms Ecology - interrelationships and problems Microbiology Laboratory procedures and research/inquiry
Human Anatomy (11) and Human Physiology (11)
Cardiovascular Endocrine Gastrointestinal Integumentary Lymphatic/immune Musculoskeletal Nervous (including exocrine) Reproductive Respiratory Urinary Senses
Chemistry (5)
Atomic Structure – includes isotopes, ions, elements, compounds, and the Periodic Table Bonding – includes Octet Rule, Ionic, Covalent, and Polar Covalent Bonds States of Matter (Gas Theory) Mixtures - Solutions, Tinctures, and Emulsions Chemical Reactions – includes physical changes; combination, single replacement, double displacement, neutralization reactions (acid/base); pH scale; catalysts; and chemical equilibrium
Health (8)
Health Dimensions – includes the meaning of health, and five dimensions of health (physical, mental, social, emotional, and spiritual) Levels of Health - individual, community, population Influencing Factors - includes exercise, hygiene, nutrition, sleep, and substance use Health Threats - substance abuse disorder, environment, health inequities Prevention and Safety - includes three levels of prevention (primary, secondary, tertiary), home safety, emergency first aid, food safety Health Screenings Vaccinations - including active, passive, and community immunity

5. Psychometric Quality

In the educational measurement field, two primary criteria are used to define high-quality information that supports the use and interpretation of a set of test scores: validity and reliability (Messick, 1989). These two criteria constitute the first two chapters of the *Standards for Educational and Psychological Testing* (APA, AERA, & NCME, 2014). Validity is concerned with the appropriateness, meaningfulness, and accuracy with which inferences can be made about a

measured characteristic; reliability addresses the consistency of measures under different conditions (APA, AERA, & NCME, 2014; Messick, 1989). This section presents estimates of reliability of NEX scores from the 2022 operational data and underscores past and future plans for collecting evidence of validity regarding the appropriateness, meaningfulness, and usefulness of inferences made about students based on their scores.

5.1 Validity

As described in the *Standards for Educational and Psychological Testing*, validity refers to the degree to which evidence and theory support the interpretations of test scores entailed by the proposed uses of tests (APA, AERA, & NCME, 2014). The NLN has a long history of pursuing construct validity, the meaning of a measure, to support test interpretation and justify test use. Additionally, the NLN has conducted validation studies to determine the usefulness of their admissions tests as predictors of student performance in schools of nursing. It is well accepted that collecting validity evidence and conducting studies is an ongoing endeavor. The NEX, as an evolution from the NLN Pre-Admission Exam (PAX), requires new validity studies which take time for data availability (e.g., student performance metrics). Nevertheless, many of the foundations upon which the PAX was developed hold steadfast and will be reiterated in this section as applicable.

Decisions about the types of evidence needed to support a validation argument depend on the claims made for how test scores may be interpreted; this can be facilitated by developing propositions (APA, AERA, & NCME, 2014). As such, NEX articulates four propositions to guide the current and future collection of validity evidence. Propositions 1 and 2 support collecting evidence for construct validity while Propositions 3 and 4 relate to criterion-related validity studies. The four propositions for the NEX are:

1. Fundamental academic abilities in math, science, and verbal knowledge and comprehension are prerequisite to success in a nursing program.
2. Test content is consistent with those prerequisite academic abilities.
3. Success in a nursing program can be assessed.
4. Students scoring high on the test will be more successful in a nursing program than students scoring low on the test.

The following sections review each proposition and how it supports or guides the collection of validity evidence, including historic, current, and future efforts.

5.1.1. Proposition 1: Academic Abilities are Pre-requisites to Success in a Nursing Program

The NEX is designed to measure abilities in verbal, math, and science as one component in the total evaluation of each applicant for both admission and support for student success decisions. Valid

use and interpretation of scores, in part, depends on the degree to which the test content represents academic abilities required for students to be successful in a nursing program. The PAX was originally built on a framework of content that SMEs identified as having importance towards success in a nursing program. During 2022 - 2023, NLN leadership committees, communities of interest, nurse faculty, and general education SMEs, reviewed the framework and validated content that should remain as well as recommended changes that would better reflect the pre-requisite abilities needed for the success of students in a nursing program. These recommendations are reflected in the blueprints and content outlines for each subject area and collectively help build evidence toward construct validity and support for Proposition 1.

Table 5 and Table 6 include clear rationales to help students and nursing educators clearly connect the pre-requisite skills to success in the nursing program for verbal ability and mathematical content. Future efforts to enhance this proposition include developing a parallel table for science content.

5.1.2 Proposition 2: Test Content Aligned to Pre-requisite Abilities

To continue building evidence for validity, the test blueprint must represent the pre-requisite abilities needed for success in a nursing program. Each test section blueprint provides the specific intersections of content, cognitive skills, and item difficulty levels to achieve a test representative of the pre-requisite knowledge, skills, and abilities needed for students to be successful in a nursing program. The test blueprints undergo review processes that include content SMEs, NLN leadership communities, and external consultants.

Test section blueprints guide item development assignments to SMEs, ensuring the content coverage requirements are met. Content SMEs are carefully selected and trained to write items that align with the blueprint; items submitted by SMEs undergo additional review and pretesting to verify content accuracy.

Propositions 5.1.1 and 5.1.2 help build evidence for construct validity. To summarize the general process[3], the NLN first builds on its long history of excellence using foundational content frameworks and engages expert reviewers to identify changes in content that intersect with the pre-requisite abilities needed for student success in a nursing program. Second, detailed blueprints and content outlines are developed to provide necessary information to successfully build a test reflective of the identified pre-requisite abilities. Third, item writers receive training about the intended population, test blueprint, content outlines, and receive item writing assignments aligning to the blueprints. And fourth, a robust review process verifies quality and accuracy is met. The formal test development process was described in Section 3 with a graphic representation provided in Appendix B.

[3] Standard 1.11 When the rationale for test score interpretation for a given use rests in part on the appropriateness of test content, the procedures followed in specifying and generating test content should be described and justified with reference to the intended population to be tested and the construct the test is intended to measure or the domain it is intended to represent.

5.1.3 Proposition 3: Defining Success Metrics in Nursing Programs

Each nursing program identifies student success metrics applicable to their program goals. Often, GPA, persistence across semesters, year-to-year retention, clinical performance, graduation, and NCLEX passing scores are indicators of success. Without the ability to define success, criterion-related validity studies cannot be conducted. Programs are encouraged to identify their success metrics and plan for validity studies, individually or collaboratively with other institutions or the NLN. With a new exam, these studies can be accomplished when performance data are available, typically after year one.

5.1.4 Proposition 4: Evidence of Relationship with Student Performance

Over the years, historic criterion-related validity studies have been conducted for the PAX-RN™ admissions exam. These studies included substantial sample sizes and participation from a variety of nursing programs and schools (NLN Pre-Admission Examination-RN™ Technical Manual, 2002). Although the degree to which the studies continue to represent the current PAX is neither ascertained nor purported in this manual, two of the studies are referenced in Table 10 to represent the solid foundation upon which the NLN built the original PAX exam, which serves as a precursor to the NEX. Future validity studies for NEX would benefit by repeating the type of work conducted by Weatherby (2007), accounting for success metrics relevant across institutions, and studies by individual institutions analyzing the predictive validity of the exam with their student population and key success metrics.

Table 10. Examples of Prior Predictive Validity Studies

Author (Year) **Scope**	**Purpose**	**Key Results**
Weatherby (2007) Multiple programs from a variety of institutions across the states	One- and two-year program retention: predictive validity	Significant mean score differences in all section tests and composite scores for students who did or did not complete their first year of program (t-test). Significant mean score differences in all section tests and composite scores for students who did or did not complete their first year of program (t-test).
Kirking (2004, cited in NLN, 2016) A single institution	Student performance in select courses over 4 semesters: predictive validity	Significant correlations between student performance and PAX composite scores. Step-wise Regression indicated composite score as the best single predictor of success in Anatomy and Physiology

5.2 Reliability

Reliability refers to quantifying the precision of scores from a measure and the extent to which scores from a given measure would be expected to vary across replications of the administrations of that measure (Haertel, 2006). Reliability estimates for NEX were computed using a base form comprised of NEX items for 71 exam sittings administered in 2023 from November 27 to December 5. The base form represents one possible rendition from the existing PAX exam which is delivered as a modified Random Parallel Test; using the base form was a reasonable means to compute NEX test statistics, shown in Table 11, while data continue to be collected.

Table 11. NEX Test Statistics

Subject	Test Length	Cronbach's Alpha	Standard Deviation	Standard Error of Measurement
Verbal	50	0.896	10.40	3.36
Math	40	0.926	9.80	2.67
Science	55	0.915	10.88	3.17
Composite	145	0.964	28.29	5.34

6. Scales, Norms, and Linking

NEX scores are reported as percentile ranks estimated from an operational normative group. This reporting facilitates the interpretation of each subject test score, Verbal Ability, Mathematics, and Science, as well as the total, or composite score. Examinees and institutions interpreting the scores can easily compare individual or cohort performance against the operational reference population in the subject areas and total performance on the exam. The NEX is comprised of a subset of the PAX items, affording the ability to be mathematically linked through common items and ability estimates computed by a Rasch model based on Item Response Theory. As mentioned in Section 3.5, this exam employs a modified Linear-On-The-Fly-Test (LOFT) delivery. The methodology for scaling, norming, and linking are described below.

6.1 Scaling and Percentile Ranks

Percentile ranks are a non-linear score transformation that sets scores on a common scale, from 1-99, and promotes score interpretation by comparing an individual's raw score to the scores of a norm

reference group. The NEX uses percentile ranks to report scores for each section and combines all section percentile rank scores, equally weighted, to create a composite score. The composite score ranges from 3 – 297; for ease of user interpretability, the scale can be referred to as 0 – 300. For example, if an examinee's Verbal, Mathematics, and Science percentile ranks were 55, 42, and 76, the composite score computes as 55 + 42 + 76 = 173. The composite score is transformed to a percentile rank and interpreted in comparison to the norming group's composite scores. Appendix C contains the transformation between raw score, percent correct, and percentile rank for the composite score and each subject test.

This scoring method provides ease of interpreting the scores by facilitating insight into the following:

- an examinee's subject level performance relative to the target population (2022 user norming group, N = 10,877),
- an examinee's overall test performance relative to the target population,
    - the higher the percentile, the more academically prepared the student is in comparison to the target population,
- an examinee's relative strengths and weaknesses across subjects
    - since all subject tests are reported on the same scale, examinees and faculty coaches can compare performance between exams
- a cohort's relative strengths and weaknesses across subjects

The use of percentile ranks alone does not designate set performance standards, levels of mastery, or a cut-score; it is distinctly a comparison of examinees to their peers.

6.2 Norms

Operational, or user norms, are based on examinees taking a test over a given time (Kolen, 2006). The advantages of using operational norms include relevance of the exam setting and context, applicability of scores from the population of interest, ease of implementation, and, in accordance with the *Standards,* the ability to update norms with a frequency that facilitates accurate and appropriate score interpretations[4]. A common disadvantage to using operational norms is a potential lack of diversity compared to a national norming group (Kolen, 2006). Disadvantages are primarily derived from the constraints applied when selecting a user group, such as limited time or locations.

The NEX uses an operational norming group of PAX examinees (N=10,877) and overcomes the difficulties in achieving a nationally representative sample through an expansive period of data collection (January 1, 2022 through December 31, 2022) without location or administration constraints.

[4] Standard 5.11 - If a test publisher provides norms for use in test score interpretation, then as long as the test remains in print, it is the test publisher's responsibility to renorm the test with sufficient frequency to permit continued accurate and appropriate score interpretations.

6.2.1 Methodology for Norming

Each participant from the 2022 operational norming group had their responses to questions in the three subject tests linked together, and a Rasch model was fit to calculate an 'ability' estimate per participant. A Rasch model estimates a set of parameters of difficulty of questions, and ability of participants, which maximize the fit between a set of observed item responses and probabilities such that:

$$P(X_{ij} = 1) = \frac{\exp(\theta_i - \delta_j)}{1 + \exp(\theta_i - \delta_j)}$$

For a given question *j*, and a given participant *i*, the probability *P* of their response *X*, being awarded a 1 relates to the difference in ability (theta, θ) of the participant and difficulty (delta, δ) of the question. A difficult question is one that is less likely to receive a correct response; a high ability participant is one that is more likely to respond correctly.

Rasch models are 'measurement models' only and make no assumptions about the causes of difficulty or ability. If a set of parameters show adequate fit to a set of observed responses, then we can predict the responses that were not observed, such as if participants were administered a different set of questions. Rasch models are therefore used widely for 'equating' or linking scores from different forms of a test, since they can take into account the difficulty of questions that different participants are administered.

The estimated model showed adequate fit to the data, with no question or participant showing infit or outfit mean-square greater than 2 or below 0.4. Questions or participants that are out of these ranges are conventionally investigated for whether the analysis would be improved by excluding these, for example if a participant is not motivated they might answer at random (where the probability of correct response is not predicted by their ability), or a question does not discriminate between high and low ability participants. The adequate fit, while not a full factor analysis, suggests that a uni-dimensional solution to understand response to questions is supported by the operational data. If the model demonstrated problematically low fit, it would be evidence that the composite score is not a meaningful measure.

Having estimated the abilities and difficulties that predict the observed data the best, a simulation of each participant sitting 100 times for each NEX subject was conducted. Each test represented a random draw of the questions according to the specification, such as a participant might be administered. Each simulated assessment had their predicted question scores summed to represent the 'raw score' that this participant would receive for each subject. Using the cumulative distribution function each raw score point was expressed as the percentile rank of the predicted data.

By this method, each percentile rank from NEX subjects reflects the proportion of the norm group of abilities from participants in calendar 2022, that we would expect to respond correctly to *this number of questions or fewer* on a randomly chosen NEX set of questions. The composite percentile rank reflects the proportion of the norm group of abilities as they would be expected to score across

the three subjects. Raw score and percentile rankings at the subject and composite score levels are shared in Appendix C.

Robust norming group data represent all examinees in the target population, inclusive of onsite and remote test delivery methods. Differences in scores between delivery methods (onsite or virtual live proctor vs. record-and-review proctor) were detected. The differences prompted an update to test administration specifications, requiring onsite or virtual live proctored events (i.e., eliminating record and review proctoring).

6.2.1 Future Operational Norms

NEX includes a pre-exam survey to facilitate the collection of demographic information including RN/PN program designation which is currently unavailable in the 2022 norming group. Data collected from the survey will support future validity studies and a prioritized analysis of RN/PN scores to determine if distinct norm reference groups are required for the most accurate and appropriate score interpretation[5].

6.3 Linking

From a practical standpoint, there is a transition period when institutions will have scores from nursing program applicants who have taken either the PAX or NEX and will need a temporary method to compare students' scores. Rescaling and changes to test specifications and administration methods can make it difficult to compare scores, resulting in tests that are not strictly equivalent[6], such as the PAX and NEX. However, as indicated by Kolen (2006) in the fourth edition of *Educational Measurement*, studies linking the scales can be helpful during the transition. Based upon a request from the nursing community for a tool to help in comparing scores across tests while continuing to maintain intended fairness to applicants, steps were taken to link data from the two exams. Results of this linkage are presented in Appendix D which includes the percentile ranks based on 2022 data as well raw scores for both the PAX and NEX exams.

6.3.1 Linking Methodology

PAX percentile ranks have historically been reported based on norm groups from 2016, RN (N=103) and PN (N=53), whereas NEX percentile ranks, as discussed in Section 6.2, are based on a 2022 operational norm group (N = 10,887). With the Rasch model as described earlier, the distribution of

[5] Standard 5.10 - When norms are used to characterize examinee groups, the statistics used to summarize each group's performance and the norms to which those statistics are referred should be defined clearly and should support the intended use or interpretation.
[6] Standard 5.20 - If test specifications are changed from one version of a test to a subsequent version, such changes should be identified, and an indication should be given that converted scores for the two versions may not be strictly equivalent, even when statistical procedures have been used to link scores from the different versions. When substantial changes in test specifications occur, scores should be reported on a new scale, or a clear statement should be provided to alert users that the scores are not directly comparable with those on earlier versions of the test.

abilities for the 2016 norm groups used in PAX were compared with the operational group that would form the 2022 norm groups used for NEX. This meant the question responses that the 2016 groups would make to the simulated NEX questions could be predicted.

A simulation of each participant from the 2016 norm group sitting 100 times for each NEX subject was conducted. Each test represented a random draw of the questions according to the specification, such as a participant might be administered. Each simulated assessment had their predicted question scores summed to represent the 'raw score' that this participant would receive for each subject. Using the cumulative distribution function, each raw score point was expressed as the percentile rank of the predicted data.

The resultant percentile ranks therefore reflect the proportion of the 2016 norm group of abilities as they would be expected to respond correctly to *this number of questions or fewer* on a randomly chosen NEX set of questions. The composite percentile rank reflects the proportion of the 2016 norm group of abilities as they would be expected to score across the three subjects.

6.3.2 Linking Considerations

Since 2016, students' overall performance on the PAX has continued to improve. This may be due to greater familiarity with the assessment, more targeted preparation and practice, or some other factor. From the methods described above, a concordance table was created to reflect the 2022 group of performances so that test takers are compared to their peers, rather than those from a historical cohort. To further aid in this transition, the NLN will report NEX percentile ranks based on the 2022 operational norming group and PAX percentile ranks based on both the 2016 and 2022 norming groups, the latter reflecting expected performance on the NEX set of questions (Reference Section 7.3, Institutuional Report).

Data from three sources (summary of test specification changes (Section 2.4), linked scores (Appendix D), and noted improvement in student performance with the 2022 norms) provide nursing faculties and administrators with sufficient evidence for decision making. Furthermore, these data enable faculties and administrators to consider their unique student populations and the degree to which the testing differences between the PAX and NEX might affect students' exam performance as well as develop an effective plan for transition that promotes fairness to their applicants.

7. Reporting and Quality Controls

There are three primary scoring reports[7] available for the NEX exam: (1) The Participant Report, available to the examinee and the school administrator to review a single examinee's scores; (2) The Institutional Coaching Report, available to the examinee and school administrator

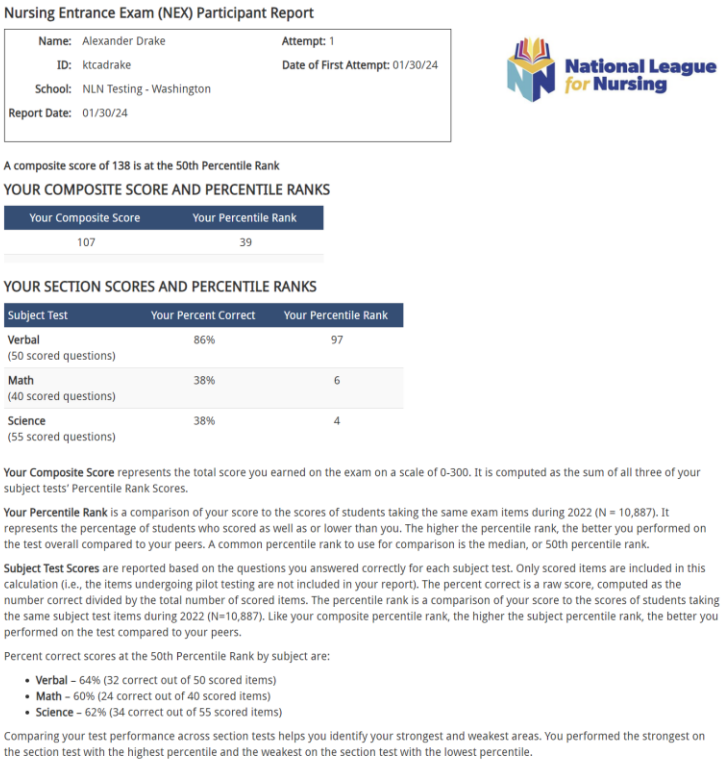
[7] Standard 6.10 - When test score information is released, those responsible for testing programs should provide interpretations appropriate to the audience. The interpretations should describe in simple language what the test covers, what scores represent, the precision/reliability of the scores, and how scores are intended to be used.

to review a single examinee's scores for the purpose of insights into performance and guide remediation if needed; and (3) The Institutional Report, intended for the school administrator to aggregate all applicants scores in one report. This technical manual serves as the source for precision and reliability of scores in lieu of adding those statistics to the report; this decision was based on feedback from end-users indicating the statistical elements create more confusion for students and administrators when interpreting the score results. The information for each report is provided below, followed by a section addressing quality control processes.

7.1 Participant Report

Each examinee and approved administrator(s) at institutions listed by the examinee, have access to a copy of the Participant Report which displays an examinee's composite score and percentile rank based on the 2022 operational normed-user group (Section 6.2). It also provides the percent correct score and percentile rank for each subject test. The report is intended to provide a non-complex view of an examinee's scores and support the interpretation of those scores by including context and definitions within the report in simple language. Using the subject test percentile ranks, a student can determine their relative strengths and weaknesses between assessed subjects. Figure 1 depicts a sample Participant Report and the verbiage included to help students interpret the scores.

Figure 1. NEX Participant Report Sample



Nursing Entrance Exam (NEX) Participant Report

Name:	Alexander Drake	**Attempt:**	1
ID:	ktcadrake	**Date of First Attempt:**	01/30/24
School:	NLN Testing - Washington		
Report Date:	01/30/24		

National League for Nursing

A composite score of 138 is at the 50th Percentile Rank

YOUR COMPOSITE SCORE AND PERCENTILE RANKS

Your Composite Score	Your Percentile Rank
107	39

YOUR SECTION SCORES AND PERCENTILE RANKS

Subject Test	Your Percent Correct	Your Percentile Rank
Verbal (50 scored questions)	86%	97
Math (40 scored questions)	38%	6
Science (55 scored questions)	38%	4

Your Composite Score represents the total score you earned on the exam on a scale of 0-300. It is computed as the sum of all three of your subject tests' Percentile Rank Scores.

Your Percentile Rank is a comparison of your score to the scores of students taking the same exam items during 2022 (N = 10,887). It represents the percentage of students who scored as well as or lower than you. The higher the percentile rank, the better you performed on the test overall compared to your peers. A common percentile rank to use for comparison is the median, or 50th percentile rank.

Subject Test Scores are reported based on the questions you answered correctly for each subject test. Only scored items are included in this calculation (i.e., the items undergoing pilot testing are not included in your report). The percent correct is a raw score, computed as the number correct divided by the total number of scored items. The percentile rank is a comparison of your score to the scores of students taking the same subject test items during 2022 (N=10,887). Like your composite percentile rank, the higher the subject percentile rank, the better you performed on the test compared to your peers.

Percent correct scores at the 50th Percentile Rank by subject are:

- **Verbal** – 64% (32 correct out of 50 scored items)
- **Math** – 60% (24 correct out of 40 scored items)
- **Science** – 62% (34 correct out of 55 scored items)

Comparing your test performance across section tests helps you identify your strongest and weakest areas. You performed the strongest on the section test with the highest percentile and the weakest on the section test with the lowest percentile.

7.2 Institutional Coaching Report

Each examinee and approved administrator(s) at institutions listed by the examinee, have access to a copy of the Coaching Report. The report provides the percentage correct for a given subject test and for each of the content areas (topics) within the subject test (e.g., the math test will report a total math test percentage correct and the percentage correct for each of the content areas: algebra, measurement, data & information, and numbers & operations). Additionally, an average percentage correct for the total test and content areas is provided and represents the average percentage correct based on all examinees who have the test at the time the report is generated. The average continues to refresh as more examinees complete the test. The NEX, as a new test, will need time to collect data before the averages are stable. The average percentage correct (mean) should not be confused with the 50th Percentile (median) as those are two different statistics generated from two different populations, all examinees who have taken the test thus far, and the 2022 operational group, respectively. The information provided in the Institutional Coaching Report can be used as a guide for students admitted to a nursing program to know their strengths or weaknesses and prepare accordingly, or for remediation purposes for those who opt to retake the exam to achieve a higher score. Figure 2 depicts a sample of a Institutional Coaching Report for the Math test.

Figure 2. NEX Institutional Coaching Report Sample

National League for Nursing

NLN Student Score Report

Participant first name	Alexander	Date/time finished	Jan 30 2024 15:36:03
Participant last name	Drake	Status	Finished
Participant ID	ktcadrake	Total score	15
Participant group	NLN-01 - NLN Testing - Washington	Percentage score	38%
Assessment name	NEX MATH	Time taken	00:03:38

Comparisons

Description		Your Scores
Number of results in comparison set: 7		
NEX		
Comparison	**%**	
▼ Participant	38%	
▲ Avg score	41%	
NEX\Math\Algebra		
Comparison	**%**	
▼ Participant	43%	
▲ Avg score	37%	
NEX\Math\Data and Information		
Comparison	**%**	
▼ Participant	33%	
▲ Avg score	50%	
NEX\Math\Measurement		
Comparison	**%**	
▼ Participant	36%	
▲ Avg score	38%	
NEX\Math\Numbers and Operations		
Comparison	**%**	
▼ Participant	42%	
▲ Avg score	43%	

7.3 Institutional Report

An institution's approved administrator(s) can create a report that lists scores for all examinees applying to the institution's nursing program, selecting date ranges to create cohorts of applicants as desired. Examinees must list the institution as a recipient of their score to be included in the report. The report displays examinees' composite scores and associated

percentile ranks, and subject test percentile ranks. This report supports the institution's review of a collection of potential applicants and how they rank in comparison to the 2022 user norming group. Subject test percentile ranks can be used to plan for broad subject areas resourcing for an applicant or cohort of applicants' success in a nursing program. Figure 3 depicts a sample of an Institutional Report, and interpretative comments included in the report, for one applicant; additional applicants' information continues as rows are added to the report.

Figure 3. Institutional Report Sample

QM Academy - Miami

Date Created: 02/02/2024

National League for Nursing

Nurse Entrance Exams: List of Students and Scores

					Composite			Verbal (50 Questions)		Math (40 Questions)		Science (55 Questions)	
Student Name	ID	Exam	Date	Attempt Number	Score	Percentile Rank	Other Reference	Percentile Rank	Other Reference	Percentile Rank	Other Reference	Percentile Rank	Other Reference
Eisenberg, Howard	eisenberg.h@icloud.com	Nursing Entrance Exam (NEX)	01/30/24	1	129	46	n/a	97	n/a	13	n/a	19	n/a

Date is when the student completed the exam.

Attempt Number:

Institutions are responsible for determining the number of attempts allowed per student and enforcing the amount of time between attempts. If a student has a total of 3 attempts, twice during the report date selection range, and once before the date selection range, attempts 2 and 3 will appear in this report.

Composite Score represents the total score earned on the exam on a scale of 3-297. It is computed as the sum of all three subject tests' Percentile Ranks.

Percentile Rank is a comparison of a student's score (composite or subject) to scores of students taking the same exam items during 2022 (N = 10,887). It represents the percentage of students who scored as well or lower than a given percentile rank. A common percentile rank to use for comparison is the median, or 50th percentile rank.

A composite score of 138 is at the 50th Percentile Rank

Percent correct scores at the 50th Percentile Rank by subject are:

- **Verbal** – 64% (32 correct out of 50 scored items)
- **Math** – 60% (24 correct out of 40 scored items)
- **Science** – 62% (34 correct out of 55 scored items)

Please reference the Technical Manual for additional details.

7.4 Interpreting and Using Scores

Each nursing program is encouraged to use the scores from the NEX exam as one method in an overall evaluation process of a potential applicant's preparedness for the academic rigors of a nursing program. Each program considers the acceptable level of performance on the exam based on their student population and available resources to support student success. The NLN does not establish a "passing" or "failing" mark on this exam, nor is it involved in any benchmark settings that an institution might conduct. Evidence for the valid use of test scores and interpretation of test scores provided in this technical manual is intended to help programs make informed decisions about how scores from the exam can be used as one factor in admissions decisions.

7.5 Quality Control for Continuous Mode Tests

The NEX is a Continuous Mode Test (CMT) which means it does not restrict test administrations to certain dates; instead, it is administered on-demand, either virtually with a live proctor, or on-site at a testing center approved by an institution's nursing program. Widely accepted quality control procedures are well-documented and available for large scale tests with set annual administrations on limited test forms, whereas the quality control procedures for a CMT are not as mature (Allalouf, Gutentag, & Baumer, 2017); hence they recommend a comprehensive automated quality control schema and framework for CMT quality and error

monitoring. At the most basic level, the authors provide recommendations for monitoring errors and quality controls during test administration and analysis. In addition to the quality control processes followed before test release, these two basic categories guide the NLN's quality control processes outlined below.

7.5.1 Quality Control During Test Administration

Allalouf et al. (2017) recommend six monitoring activities during test administration; four of these areas are addressed in Table 12, adapted to meet the NLN's organizational needs and capacities.

Table 12. NLN Quality Controls During Test Administration

Category	Desired Outcome	QC check	Monitoring
Test Allocation	LOFT delivery per specifications	Allocation as planned; check sample of repeat examinees for receipt of different form to the extent available with item pool	Quarterly @ one subject test per quarter
Test Allocation	Random based order of items; random response options	Item order as planned; response options as planned	Quarterly @ one subject test per quarter
Testing Conditions	Standardized testing conditions	On-site and virtual live proctors follow standardized administration guidelines. Extreme conditions highlighted and reporting. No relationship between test conditions and scores.	Annual
Problem Resolution During Testing	No issues; if issues surface, they are solved quickly	Test delivery vendor and virtual live proctor offer active help desk during testing; report on the number of calls/tickets, types of issues, and/or time to resolve and issues while testing.	Bi-annual

7.5.2 Quality Control in Test and Item Analysis

Allalouf et al. (2017) recommend eight areas to monitor during test analysis and scoring; six of these areas are addressed in Table 13, adapted to meet the NLN's organizational needs and capacities.

Table 13. NLN Quality Controls During Test Analysis

Category	**Desired Outcome**	**Process**	**Monitoring**
Item properties	Desired psychometric properties as planned.	Conduct item analysis. Items not meeting boundaries are retired or reviewed by SMEs with item edits or replacements as needed.	Quarterly @ one subject test per quarter
Item stability	Stable item properties.	Drift analysis.	Quarterly @ one subject test per quarter
Transformation into standardized scores	Score transformation to percentile ranks is stable.	Analyze relationship between raw and standardize scores	Quarterly @ one subject test per quarter; Annual for composite
Stability of scores	Scores over time are stable.	Monitor scores over time; interpret significant deviations	Quarterly @ one subject test per quarter
Mean scores	No significant differences between scores based on administration type (onsite vs. proctor).	Analyze means for significant differences and examine as needed.	Annual
Repeat examinee's scores	Score differences are within reason.	Review scores of examinees who repeat the test, the number of attempts, and track significantly higher or lower scores.	Annual

The NLN follows the extensive review and quality control procedures as outlined to detect issues in a reasonable time. Occasionally an error may be detected after scores have been reported. In the event of such a finding, scores may be recalculated, and reports revised for those students whose score would change.

8. Fairness and Equity in Testing

Following APA, AERA, and NCME, 2014, the NLN seeks to ensure that all steps in the testing process are designed to minimize construct-irrelevance and promote valid score interpretations for all examinees in the target population. Section 3 of this manual delineates processes in the test development steps to support those efforts.

NLN Assessment Testing services recognizes the diversity of applicants to nursing schools and is astutely aware of the need for fairness in testing. To maximize access to taking the NEX exam, the NLN offers onsite and virtual (remote live proctor) testing, the latter mitigates potential transportation hardships students might otherwise encounter. Score reports are produced immediately, a method intended to reduce travel during the admissions process for onsite examinees. Fairness extends to minimizing potential differences between onsite and virtual examinees. Provisions for live proctors are implemented with the NEX; results from future analysis will be published in the annual update.

Special accommodations are available in accordance with the Americans with Disabilities Act (ADA) of 1990 to eliminate discrimination against individuals with disabilities and to adhere to the standards and protections established on behalf of those individuals[8]. Special accommodations are requested through the nursing program's institutional processes; approved accommodations for additional time or readers are forwarded to the NLN.

Student and Administrator guides are provided through the NLN Testing Services Website to facilitate standardized test administration regardless of delivery method or accommodation[9]. Student guidelines help examinees understand test content, integrity in testing, opportunities to purchase practice exams, and administration requirements for onsite or virtual testing. Administrator guidelines delineate expectations of testing centers and live proctors to achieve a standardized and consistent testing experience for all examinees during the administration of the exam.

9. Other Information Regarding NLN Test Administration

9.1 Limitations of Exams

Like any exam, the NEX is not a perfect measurement of achievement. As a general test of

[8] Standard 3.9 - Test developers and/or test users are responsible for developing and providing test accommodations, when appropriate and feasible, to remove construct-irrelevant barriers that otherwise would interfere with examinees' ability to demonstrate their standing on the target constructs.

[9] Standard 3.4 - Test takers should receive comparable treatment during the test administration and scoring process.

students' knowledge and skills across Verbal, Mathematics, and Science, the test represents a sample of knowledge in each broad subject area. Because of this, questions on a specific test may be more favorable to some examinees than to others. Errors in measurement may also be introduced by examinees. Individuals who are ill or extremely tired or nervous may not perform to the best of their ability. These limitations should not be forgotten by those reviewing students' performance on the exam.

9.2 Retake Policy

To maintain the integrity, security, and validity of the exam, the NLN does not recommend students re-test for a period of 30 days and includes this statement in the student and administrator guidelines. The recommendation is not enforceable, and the NLN relies on institutions to set and enforce their own re-test policy. Institutions should be aware of students re-taking the test consecutively, potentially falsely improving their scores, by using the test as a tool of practice or memorization.

To improve scores when re-testing, students should seek feedback about their exam performance from a nursing program administrator or faculty member and spend ample time studying areas in need of improvement. The NLN offers the NEX Prep Exam as a resource for students to prepare for their initial or re-take of this exam.

9.3 Challenge Items

Examinees may challenge any items that they feel are questionable. Once a challenge to an item is raised and reported, the item is reviewed by content specialists in the content area assessed by the item. If a problem is found, actions are taken to eliminate or minimize the influence of the problem item as necessary. In all cases, the person who challenges an item is sent a letter indicating results of the review.

9.4 Test Score Retention

Test scores are the property of NLN and are retained on file in electronic form for a period of three years after the date of examination. Duplicate copies of score reports are available upon written request by submitting the Student Duplicate Score Report available on the NLN Testing Website.

9.5 Fair Testing and Ethics Statement

The NLN believes most examinees manage themselves with integrity and are disturbed when they observe others cheating. Examinees are expected to maintain the highest standards of

conduct. Using or attempting to use unauthorized assistance, material, or study aids during an examination, will result in a disqualification of the exam and future NLN exams. The NLN Fair Testing and Ethics Statement is provided in Appendix E.

References

Allalouf, A., Gutentag, T., & Baumer, M. (2017). Quality control for scoring tests administered in continuous mode: An NCME instructional module. *Educational Measurement: Issues and Practice, 36*(1), 58-68.

American Psychological Association, American Educational Research Association, & National Council for Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Becker, K. A., & Bergstrom, B. A. (2013). Test administration models. *Practical Assessment, Research & Evaluation, 18*(14). Available online: http://pareonline.net/getvn.asp?v=18&n=14 retrieved 10/27/2023.

Ebel, R. L., & Frisbie, D. A. (1986). *Essentials of educational measurement*. Englewood Cliffs, NJ: Prentice-Hall.

Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 65–110). Lanham, Maryland: American Council on Education and Rowman & Littlefield.

Kolen, M. (2006). Scaling and norming. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 155 - 186). Lanham, Maryland: American Council on Education and Rowman & Littlefield.

Lord, F. M. (1952). The relation of the reliability of multiple-choice tests to the distribution of item difficulties. *Psychometrika, 17***,** 181–194. https://doi.org/10.1007/BF02288781

Matlock-Hetzel, S. (1997, January 23-25). Basic concepts in item and test analysis [Paper presentation]. Southwest Educational Research Association Annual Meeting, Austin, TX.

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (3rd ed., p. 13-103). New York, NY: American Council on Education and Macmillan.

National Governors Association for Best Practices, Council of Chief State School Officers. (2010). Common core state standards (mathematics). Common Core State Standards Initiative | The Standards (thecorestandards.org) retrieved 7/3/2023.

National League for Nursing Technical Services. (2007). *Pre-admission examination-RN™ Technical Manual*. Washington, DC: National League for Nursing.

National League for Nursing Technical Services. (2016). *Technical manual: Pre-admission examination*. Washington, DC: National League for Nursing.

National League for Nursing. (2020). Fair testing guidelines for nursing education. Washington, DC: National League for Nursing. https://www.nln.org/docs/default-source/uploadedfiles/advocacy-public-policy/fair-testing-guidelines.pdf?sfvrsn=f363dd0d_0 retrieved 3/14/2024.

Oermann, M. H., & Gaberson, K. B. (2021). Test and item analysis: Interpreting test results. In M. H. Oermann & K. B. Gaberson, *Evaluation and testing in nursing education* (6th ed., pp. 221-250). New York: Springer Publishing Company.

Thorndike, R. M., Cunningham, G. K., Thorndike, R. L., & Hagen, E. P. (1991). *Measurement and evaluation in psychology and education* (5th ed.). New York: MacMillan. Weatherby, F. (2007). NLN's Pre-admission examination- RN: A validation study. New York, NY: National League for Nursing.

Yoo, H., & Hambleton, R. (2019). *Foundations of operational item analysis.* National Council on Measurement in Education: Instructional Topics in Educational Measurement Series. Digital Module 08v1.0. https://360.articulate.com/review/content/0f829470-5f0d-4336-a23b-bca48e683d8a/review

Appendix A
Table of Contributors

NLN Staff

Dodie Serafini, MSN, PhD, RN, CNE	Director of Assessment Services
Karen Drake, MEd	Assessment Services Manager

Subject Matter Experts

Denise M. DeSalvo, MS, RDN, LD	The University of Alabama, Alabama
Gail Stewart, PhD	Hillsborough Co. Public Schools, Florida
Jeffery M. Schwehm, PhD	Bryan College of Health Sciences, Nebraska
Joanna B. Ambrose, MAT	West Point Public Schools, Virginia
Krisel Nagallo, BSc, BA, NMD	Chamberlain University, Arizona
Lietta Scott, PhD	Retired, Director of Psychometrics AZDOE, Arizona
Marina V. Blakely, PhD	Arizona College of Nursing, Arizona
Mark A. Jones, PhD	Bryan College of Health Sciences, Nebraska
Melissa Stinson-Borg, BS	StrongMind, Arizona
Rebecca Lovato, MAEd	Hanover County Public Schools, Virginia

Reviewers

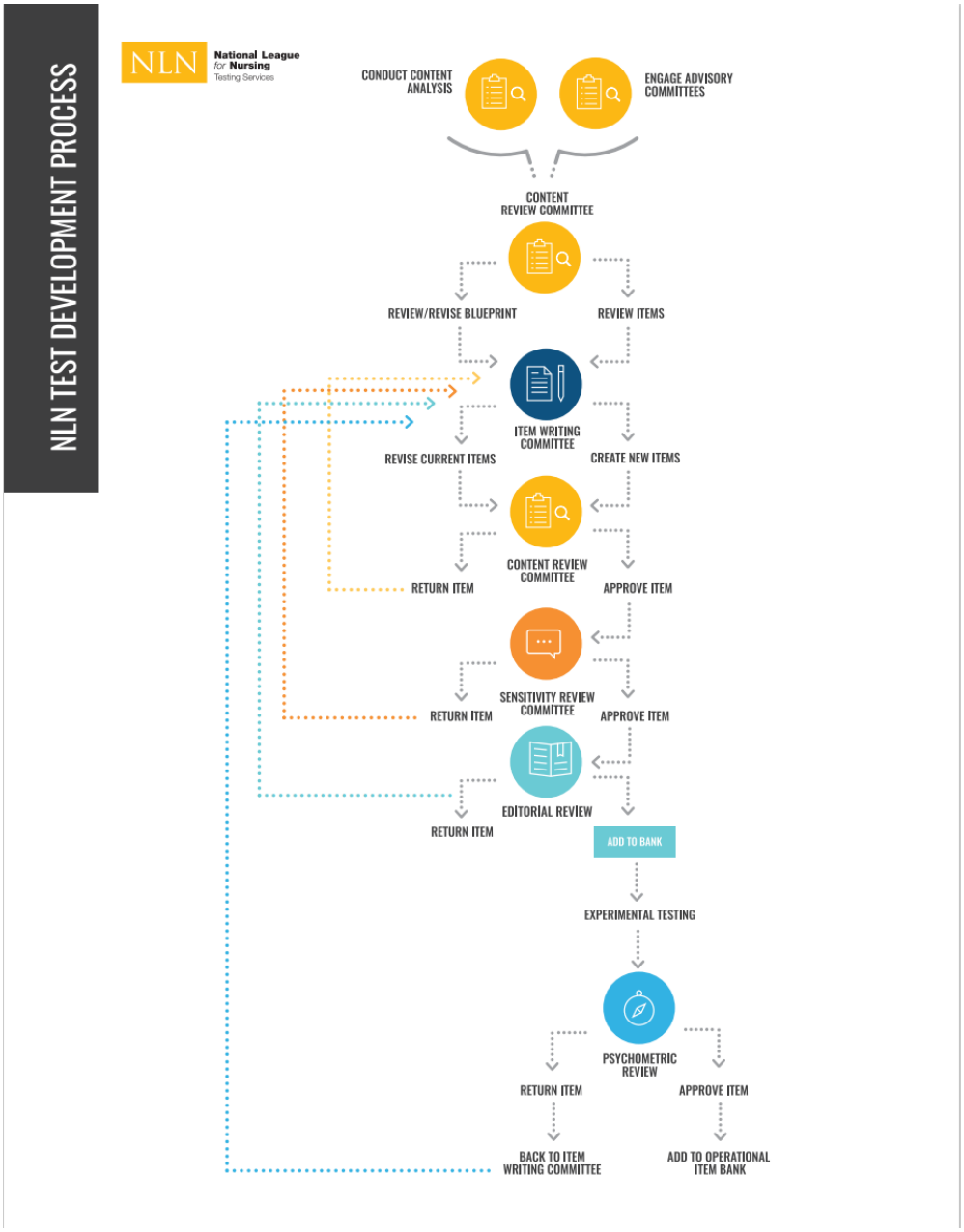
Aaron Cyr, EdD, RN	Arizona College of Nursing, Arizona
Amy Jausch, DNP, RN, CNE	Ohio State University, Ohio
Anna Jessep, DNP, APRN, FNP-C	Galen College of Nursing, Texas
Brooke Kistler, PhD, RN, CNE	Catawba Valley Community College, North Carolina
Cheryl Osler, EdD, MS, MSN, LMHC, RN, CNE	Spokane Community College, Washington
Chris Tenna-Perkins, PhD, RN, ANP, CNE	Bon Secours Memorial School of Nursing, Virginia
Katie Sladkey, EdD, MSN, RN	Bryan College of Health Sciences, Nebraska
Lisa Hawthorne, PhD, RN, PHN. CNE	Galen College of Nursing, Texas
Lisa Radesi, DNP, MSN, RN, CNS, PHN, NE-BC, CNE	Alliant International University School of Nursing and Health Sciences, Arizona
Michael Newsome, CRA	Virginia Commonwealth University, Virginia
Mollie Norgren, PhD, MSN, RN	Galen College of Nursing, Utah

Consultants

Lisa M. Mitchell, MA	Mitchell Kaizen Consulting, LLC Psychometric Evaluator
Sharron P. Schlosser, PhD, RN, CNE, COI, ANEF	Professor Emerita Samford University

Appendix B

NLN Test Development Process



Appendix C

Transformation Tables from Raw Score to NEX Percentile Ranks[10]

Table C.1. Verbal Test Score Transformation

Raw Score	Percentage Correct	Percentile Rank
0	0	1
1	2	1
2	4	1
3	6	1
4	8	1
5	10	1
6	12	1
7	14	1
8	16	1
9	18	1
10	20	1
11	22	1
12	24	1
13	26	1
14	28	1
15	30	1
16	32	1
17	34	1
18	36	1
19	38	2
20	40	3
21	42	4
22	44	6
23	46	9
24	48	12
25	50	15
26	52	19
27	54	23
28	56	28
29	58	33
30	60	39
31	62	45
32	64	51
33	66	57
34	68	63
35	70	69
36	72	74
37	74	79
38	76	84
39	78	88
40	80	91
41	82	94
42	84	96
43	86	97
44	88	99
45	90	99
46	92	99
47	94	99
48	96	99
49	98	99
50	100	99

[10] Reference Section 6.2 for norming details

Table C.2. Math Test Transformation Table

Raw Score	Percentage Correct	Percentile Rank
0	0.0	1
1	2.5	1
2	5.0	1
3	7.5	1
4	10.0	1
5	12.5	1
6	15.0	1
7	17.5	1
8	20.0	1
9	22.5	1
10	25.0	1
11	27.5	1
12	30.0	2
13	32.5	3
14	35.0	4
15	37.5	6
16	40.0	9
17	42.5	13
18	45.0	17
19	47.5	23
20	50.0	28
21	52.5	34
22	55.0	40
23	57.5	47
24	60.0	54
25	62.5	60
26	65.0	67
27	67.5	73
28	70.0	79
29	72.5	84
30	75.0	88
31	77.5	92
32	80.0	94
33	82.5	96
34	85.0	98
35	87.5	99
36	90.0	99
37	92.5	99
38	95.0	99
39	97.5	99
40	100.0	99

Table C.3. Science Test Transformation Table

Raw Score	Percentage Correct	Percentile Rank
0	0.0	1
1	1.8	1
2	3.6	1
3	5.5	1
4	7.3	1
5	9.1	1
6	10.9	1
7	12.7	1
8	14.5	1
9	16.4	1
10	18.2	1
11	20.0	1
12	21.8	1
13	23.6	1
14	25.5	1
15	27.3	1
16	29.1	1
17	30.9	1
18	32.7	2
19	34.5	2
20	36.4	3
21	38.2	4
22	40.0	6
23	41.8	8
24	43.6	10
25	45.5	13
26	47.3	16
27	49.1	19
28	50.9	23
29	52.7	27
30	54.5	31
31	56.4	36
32	58.2	40
33	60.0	45
34	61.8	50
35	63.6	55
36	65.5	60
37	67.3	65
38	69.1	70
39	70.9	75
40	72.7	79
41	74.5	83
42	76.4	86
43	78.2	89
44	80.0	92
45	81.8	94
46	83.6	96
47	85.5	97
48	87.3	98
49	89.1	99
50	90.9	99
51	92.7	99
52	94.5	99
53	96.4	99
54	98.2	99
55	100.0	99

Table C.4. Composite Score Transformation Table

Composite Score	Percentile Rank
3	1
4	2
5	2
6	3
7	3
8	3
9	4
10	4
11	4
12	5
13	6
14	6
15	6
16	6
17	7
18	7
19	8
20	8
21	9
22	9
23	9
24	9
25	9
26	10
27	10
28	10
29	12
30	12
31	12
32	13
33	13
34	13
35	13
36	13
37	13
38	15
39	15
40	15
41	16
42	16
43	16
44	17
45	17
46	17
47	17
48	18
49	19
50	19
51	19
52	19
53	20
54	20
55	20
56	22
57	22
58	22
59	22
60	22
61	23
62	23
63	23
64	23
65	24
66	24
67	24
68	24
69	24
70	27
71	27
72	27
73	27
74	28
75	28
76	28
77	28
78	28
79	29
80	29
81	29
82	29
83	29
84	31
85	31
86	31

Composite Score	Percentile Rank
87	31
88	33
89	33
90	33
91	33
92	33
93	34
94	34
95	34
96	34
97	34
98	35
99	36
100	36
101	36
102	36
103	36
104	39
105	39
106	39
107	39
108	39
109	40
110	40
111	40
112	40
113	40
114	41
115	41
116	41
117	41
118	41
119	41
120	44
121	44
122	44
123	44
124	44
125	45
126	46
127	46
128	46
129	46

Composite Score	Percentile Rank
130	46
131	47
132	47
133	47
134	47
135	47
136	47
137	47
138	50
139	50
140	50
141	50
142	50
143	51
144	51
145	51
146	51
147	51
148	51
149	54
150	54
151	54
152	54
153	54
154	54
155	55
156	55
157	55
158	55
159	55
160	57
161	57
162	57
163	57
164	57
165	57
166	60
167	60
168	60
169	60
170	60
171	60
172	60

Composite Score	Percentile Rank
173	61
174	61
175	61
176	61
177	61
178	63
179	63
180	63
181	63
182	63
183	63
184	65
185	65
186	65
187	65
188	65
189	67
190	67
191	67
192	67
193	67
194	67
195	69
196	69
197	69
198	69
199	69
200	69
201	70
202	70
203	70
204	70
205	70
206	73
207	73
208	73
209	73
210	73
211	73
212	74
213	74
214	74
215	74

Composite Score	Percentile Rank
216	74
217	75
218	75
219	75
220	75
221	75
222	75
223	78
224	78
225	78
226	78
227	78
228	79
229	79
230	79
231	79
232	80
233	80
234	80
235	80
236	80
237	83
238	83
239	83
240	83
241	83
242	83
243	83
244	83
245	83
246	84
247	84
248	84
249	84
250	84
251	86
252	86
253	86
254	86
255	88
256	88
257	88
258	88

Composite Score	Percentile Rank
259	88
260	88
261	88
262	88
263	89
264	89
265	89
266	91
267	91
268	91
269	92
270	92
271	92
272	92
273	92
274	92
275	94
276	94
277	94
278	94

Composite Score	Percentile Rank
279	94
280	94
281	94
282	95
283	96
284	96
285	96
286	96
287	96
288	96
289	97
290	97
291	97
292	98
293	98
294	99
295	99
296	99
297	99

Appendix D

PAX and NEX Conversion Table 2022 Norms

Percentile Rank 2022 Norms[11]	RN/PN PAX Score	NEX Score
1	58 or lower	3
1	59	3
2	60	4
2	61	4
2	62	4
2	63	5
3	64	6
3	65	6
3	66	7
3	67	8
4	68	9
4	69	11
5	70	12
5	71	12
6	72	13
6	73	16
7	74	17
7	75	18
8	76	19
9	77	21
10	78	26
10	79	28
11	80	29
12	81	29
13	82	32
14	83	38
15	84	38
16	85	41
17	86	44
18	87	48
19	88	49
20	89	53
22	90	56
23	91	61
24	92	65
25	93	70
27	94	70

[11] Reference Section 6.3 for linking details

Percentile Rank 2022 Norms[11]	RN/PN PAX Score	NEX Score
28	95	74
29	96	79
31	97	84
32	98	88
33	99	88
35	100	98
36	101	99
38	102	104
39	103	104
41	104	114
43	105	120
44	106	120
46	107	126
47	108	131
49	109	138
50	110	138
51	110	143
52	111	149
54	112	149
55	113	155
57	114	160
58	115	166
60	116	166
61	117	173
63	118	178
65	119	184
66	120	189
68	121	195
69	122	195
71	123	206
72	124	206
74	125	212
75	126	217
76	127	223
78	128	223
79	129	228
80	130	232
82	131	237
83	132	237
84	133	246
85	134	251
86	135	251

Percentile Rank 2022 Norms[11]	RN/PN PAX Score	NEX Score
87	136	255
88	137	255
89	138	263
90	139	266
91	140	266
91	141	268
92	142	269
93	143	275
94	144	275
94	145	281
95	146	282
95	147	282
96	148	283
96	149	288
97	150	289
97	151	291
98	152	292
98	153	292
98	154	292
98	155	293
99	156 or higher	294 or higher

Appendix E

NLN Fair Testing and Ethics Statement

NLN National League for Nursing Testing Services

FAIR TESTING AND ETHICS STATEMENT

We believe most examinees manage themselves with integrity and are disturbed when they observe others cheating. Examinees are expected to maintain the highest standards of conduct. The following are prohibited:

1. Launching a secure exam outside the proctored testing site.
2. Copying from others.
3. Providing or receiving information on all or part of an exam (e.g., telling someone or publishing what test questions appeared on your exam or being told this information)
4. Having or using a "cheat sheet" (i.e., a piece of paper with answers, formulas, information or notes) that has not been specifically authorized.
5. Having or using a communication device such as a cell phone, pager, PDA, or electronic translator to send or obtain unauthorized information.
6. Taking an exam for another student or permitting someone else to take a test for you.
7. Asking another individual to give you improper assistance, including offering money or other benefits.
8. Asking for or accepting money or any other benefit in return for giving another individual improper assistance.
9. Altering and resubmitting an assessment report.
10. Gaining or providing unauthorized access to examination materials.

Note: During an exam, the possession of any prohibited or unauthorized information or device, whether or not it is used, is an act of dishonesty. [NOTE NLN reserves the right to abort a test record, lockout a student account, and suspend the student from future testing pending an investigation and final discovery.]

Contact NLN at 800-732-8656 if you suspect any unfair or unethical testing situation.

NATIONAL LEAGUE *FOR* NURSING (NLN)
Testing Services
2408A Lebanon Avenue Shiloh, Illinois 62221 (P) 800-732-8656 (F) 202-888-3104 Rev 01/2018