

The Readability of NCLEX® Examinations

Thomas R. O'Neill, PhD
NCSBN Testing Services

The NCLEX® examinations are designed to test the knowledge and abilities essential to the safe and effective practice of entry-level nursing. Accordingly, the content of the questions is about nursing; nevertheless, all of the questions require some reading. Because the purpose of the NCLEX examinations is to measure nursing ability, not reading ability, the reading demands of the test should not be so high that the readability of the text becomes a barrier to otherwise qualified candidates. Consequently, the difficulty of an item should be governed by the nursing content rather than the semantic or syntactic complexity of the text.

To address this concern, NCSBN assesses the readability (the ease with which one can read and comprehend text) of each operational item pool before the pool is deployed for use. Readability, as used in this paper, includes the notions of semantic and syntactic complexity, but excludes aspects related to vision and perception such as screen position, screen color, font color, font size, etc. These perceptual aspects are important and were considered in the design of the test delivery software.

How is readability assessed on the NCLEX?

SELECTING AND EDITING TEXT

Before an operational item pool is deployed, the readability level of the pool is assessed. Three tests are simulated from the new pool: a minimum-length easy test, a maximum-length borderline difficulty test and a minimum-length difficult test. Because the items for these tests are from very different sections (with regard to item difficulty) of the item pool, it is unlikely that there would be overlapping items across the three tests. These items (approximately 18% of an operational pool)

are then considered as a representative sample of the items in the operational pool. The selected items are first edited to exclude all tables of information, charts and graphics. Complete sentences including technical terms, numerals, abbreviations and measurements are retained. If a sentence is spread across the item's stem and distracters, the beginning of the fragment from the stem is excluded, but the fragment is repeated with each distracter to form complete sentences. When distracters contain phrases that do not complete a sentence that began in the stem, that phrase is excluded. However, subsequent complete sentences in the distracters are included.

THE FRI

The Fry Readability Index (FRI; Fry, 1968) considers readability as a combination of sentence length and the number of syllables per word. However, NCSBN uses a variant on the procedure. The typical FRI is based upon three 100-word segments that are selected randomly from the text and the average number of syllables and the average sentence length across the three

samples is computed. These numbers are plotted on a chart to produce a grade-level readability estimate. If there is too much variability between the three samples, more samples are included and the caveat is made that the readability of the text is uneven. NCSBN uses this same procedure but, rather than select three 100-word samples, the number of syllables and sentences is counted for each simulated test (which is noticeably longer than 100 words). This count is divided by the number of words in the simulated test and multiplied by 100 to make it comparable to the 100-word frame of reference that the chart requires. The conversion chart is used to produce the grade level. The average readability of these three simulated tests is considered to be representative of the readability of the operational item pool.

LEXILES®

Lexiles® also consider readability as a function of sentence length and the word difficulty, but rather than using the number of syllables as an indicator of word difficulty, Lexiles use the frequency with which words are actually used in the written language (essentially a familiarity-rarity continuum). When improvements in computer technology made it possible to analyze enormous bodies of literature and produce frequencies for every word encountered, databases of relative word frequency were created. Once these databases existed, it was possible to use them prospectively to identify the average word difficulty for a particular text. The number of syllables in a word was a good proxy for how difficult it was to read and understand, but it did not quite match the understandability-recognition notion of word difficulty. Obscure words like furl, glib, gape and squib (one syllable) would be considered much easier than words like understanding, remembering, disagreement and hesitation (four syllables). Today the Lexile Analyzer is able to automate the process of describing the readability of a given text and produces results on an interval scale. The analyzer is used to determine the readability of each simulated test in Lexiles. The average readability of the three simulated tests is considered to be representative of the readability of the operational item pool. It is interesting to note that many publishers are now having their textbooks Lexiled so that professors have a general idea of how difficult the book will be to read. Many nursing textbooks are more difficult to read than NCLEX items.

Comparing Lexiles and FRI

The advantages of the Lexile Framework® over NCSBN's version of the FRI method are: (1) Lexiles produce a comprehension-based metric rather than a pronunciation-based metric, (2) Lexile "rulers" quickly communicate reading difficulty, are available through the Web and can be customized for special situations, and (3) the Lexile Framework is more stable because it is tied to reading material (criteria-referenced) rather than highly variable groups like school grade levels (norm-referenced). However, it should be noted that many consumers are more comfortable with the notion of grade-level even if it varies from region to region and is not metric.

What are the NCLEX standards for readability?

From 1993 through 2003, the policy regarding acceptable readability levels held that:

The readability level of the PN item pool should not exceed 8th grade reading level excluding technical terms. The readability level of the RN item pool should not exceed 10th grade reading level excluding technical terms.

However, in 2003, the policy was changed to include technical terms. The rationale for this was that candidates in the process of answering questions did have to read the technical terms so these terms should be included in the text being evaluated.

The inclusion of technical terms made the text appear more difficult to read, even when the same items were evaluated. It is interesting to note that this policy change did not revise the thresholds of acceptable readability. Such a revision could happen in the future. Presently, the policy states:

The readability level of the PN item pool should not exceed 8th grade reading level (with a corresponding range on the Lexile scale). The readability level of the RN item pool should not exceed 10th grade reading level (with a corresponding range on the Lexile scale).

Advice for educators writing practice questions

Nurse educators should not feel that they must adhere to these standards when they write questions for their students. It is perfectly fine to write items that require greater reading ability, provided that one is confident that the students reading the questions will understand what the question is asking. In fact, if students are accustomed to receiving questions that demand a higher reading ability, then the reading demands of an NCLEX examination should present no problems at all.

NCSBN assumes that all NCLEX candidates can read text presented on a computer screen and on paper. If a candidate cannot, then it is the candidate's responsibility to request an appropriate ADA accommodation from the board of nursing. The board of nursing makes the decision to approve or reject the request, as well as to determine what impact the disability has upon the candidate's fitness to practice.

To learn more about readability, the following articles and Web sites are suggested:

Fry, E. (1968). A readability formula that saves time. *Journal of Reading*, 11 (7), 265-71.

McLaughlin, G. (1969). SMOG grading: A new readability formula. *Journal of Reading*, 12 (8) 639-646.

<http://www.lexile.com/EntrancePageFlash.html>.

<http://www.timetabler.com/reading.html>.

<http://school.discovery.com/schrockguide/fry/fry.html>.



National Council of State Boards of Nursing, Inc.
111 E. Wacker Drive, Suite 2900, Chicago, IL 60601-4277
312.525.3600 | International Calls: +1.312.525.3688
Testing Services: 312.525.3750 | www.ncsbn.org