Summary

• The NCLEX passing standard changed in April 2013.
• As the passing standard has changed (increased) this means that a student must be at a higher ability level to achieve the passing standard.
• We examined the relationship between difficulty level in PrepU and IRT analysis difficulty.
• Results of the analysis revealed a significant, positive correlation between the IRT-calculated question difficulty and the prepU-determined question difficulty.
• Conclusion: a student who is consistently answering questions in prepU (on a breadth of client needs/topics) and has achieved a ML of 4 and above on all topics, is answering questions at the same difficulty level as those students who achieve the passing.

Adaptive Testing

In contrast to classical testing models, systems which incorporate features of adaptive testing help focus student learning on the just the right content in a more efficient way. Within the adaptive quizzing area of NCLEX-RN 10,000, a student’s ability level is determined and continuously updated based on their responses to calibrated items with known difficulty parameters. As students answer questions in NCLEX-RN 10,000, they achieve a Mastery Level (ML) on topics on which they have taken quizzes. A ML is essentially a measure of the average difficulty level of the questions a student answers correctly. As a student is able to answer increasingly difficult questions correctly, s/he is given even more challenging questions on future quizzes. And as these questions are answered correctly—reflecting a student’s greater and greater mastery of course concepts, the student moves up in Mastery Level.

For students using NCLEX-RN 10,000 to help prepare for the NCLEX, it is helpful to have some idea about how ML is related to the likelihood they will pass the NCLEX exam. But the question of how to relate the mastery levels in NCLEX-RN 10,000 (which are a reflection of question difficulty as described above) and the likelihood of passing the NCLEX is not easily addressed. By using Item Response Theory (IRT) however, we were able to use information from the NCSBN relating to passing standards, to help us make some educated predictions.

NCSBN Passing Standard

The NCLEX-RN exam is a variable length computerized adaptive test (CAT) and can be anywhere from 75 to 265 items long. The length of the exam is determined by the candidate’s responses to the questions. After reaching a minimum of 75 questions testing stops when the candidate’s ability is determined (at 95% certainty) to be either above or below the passing standard. The passing standard is defined as “A cut point along an ability range that marks the minimum ability level requirement. For the
NCLEX, it is the minimum ability required to safely and effectively practice nursing at the entry-level," (NCBSN). The NCSBN Board of Directors re-evaluates the NCLEX passing standard every three years when the test plan is reviewed. In 2009 the NCLEX-RN passing standard was changed from -0.21 to -0.16 logits, and then in April, 2013 the passing standard was raised again to 0.00 logits.

Logits and IRT Models

Adaptive testing, sometimes called Computerized Adaptive testing or CAT was born from the desire to create tests which could give better and more precise determinations of student ability and performance in an efficient way. A conventional test is assembled to serve the needs of an average student, and includes a set of items from a homogenous pool of items. In an adaptive test, each of the examinees is given an individually tailored sequence of items from a pool of items which differ in difficulty and reliability.

In adaptive testing we want to give each person items that will be the most informative and that will allow us to make the most precise measurement of ability given the amount of time given for the testing. Item Response Theory (IRT) is a model that allows for an examinee A, to be shown a particular item with regard to parameters for that examinee and that item. IRT helps us describe what happens when an examinee sees a test item. We want the items to measure the same construct. IRT formalizes this and defines a scale for the underlying variable (typically called proficiency) which is measured by the test items. This feature of IRT means that we can compare students even if they do not answer the same items (Thissen & Orlando, 2001).

The IRT model gives a probability of answering a question correctly in terms of the interaction between difficulty and proficiency (both of which we cannot see). Given that ability is a “hypothetical construct” we must measure it by an analogy with an observable variable (Thissen & Orlando, 2009). And so a simple IRT model determines the probability that someone with a proficiency of θ will respond correctly to an item of difficulty level b.

Difficulty (b) = the position that an item occupies on this dimension.

Proficiency (θ) = is the position of each examinee on this dimension.

The logit is the way we can measure the relative differences between our estimates of candidate ability and difficulty of the test questions. Logits are typically reported from -3 to +3, although they are infinite in either direction. The graph below gives information on how well students of different ability levels performed on different test items. Notice in the graph that the line (function) increases continually. This means that for these items, the probability of getting an item correct increases with the proficiency level of the examinee. Which is what we would expect.

As you can see on the graph (below), the higher the value of the estimate of ability, θ, the greater the ability (in whatever you are measuring) the candidate has. Similarly, the higher the estimate of difficulty, b, the more difficult the test item is and the greater chance there is to answer the question incorrectly. And so as the proficiency increases, so does the probability of answering an item correctly.

In the NCLEX example, the ability estimate (in logits) is related to the passing standard set by the NCSBN. The ability (θ) increases as the logit value increases. And so as the passing standard has changed (increased) this means a student must be at a higher ability level to achieve the passing standard.

NCLEX-RN 10,000 Item Difficulty and IRT

In order to determine the relationship between question difficulty level in the NCLEX-RN 10,000 and IRT analysis difficulty we calculated Spearman rank correlations between the IRT-determined difficulties and the question difficulties previously
Comparing IRT Difficulty and NCLEX-RN 10,000 question difficulty ratings for the top 100 items and 2,000 student responses

determined by the NCLEX-RN 10,000. This analysis used the top 100 NCLEX-10,000 items (in terms of number of times answered) and 2,000 students’ worth of data.

Results of the analysis revealed a significant, positive correlation between the IRT calculated difficulty and the NCLEX-RN 10,000 question difficulty, $r(98) = .638, p < .0005$ (see Table 1).

Table 1: Correlation between IRT b parameter and NCLEX-RN 10,000 difficulty.

<table>
<thead>
<tr>
<th>NCLEX-RN 10,000 Diff</th>
<th>PrepU Diff</th>
<th>IRT Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correlation Coefficient</td>
<td>1,000</td>
<td>0.638**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

**Correlation is significant at the 0.01 level (2-tailed).**

Based on this analysis, a student who is at 0.00 logits in terms of ability is correctly answering questions equivalent to a difficulty level of about 25 +/- 5 on the NCLEX-RN 10,000 difficulty scale. As mentioned earlier, a mastery level is essentially a measure of the average difficulty level of the questions a student answers correctly. As a student answers more difficult questions correctly, s/he is given even more challenging questions on future quizzes. If these questions are answered correctly—allowing the student to demonstrate greater and greater mastery on course concepts, the student moves up in Mastery Level.

PrepU NCLEX-RN 10,000 questions with a difficulty of 25 +/- 5 are answered by students in the ML range between 3-5. The change in passing standard from -0.16 logits to 0.00 logits did not alter the ML range for students with the associated ability level (as can be seen in Figure 1).

Although we cannot predict with 100% certainty how a student will perform on the NCLEX exam, what we can say is this: a student who is consistently answering questions in NCLEX-RN 10,000 (on a breadth of client needs/topics) and has achieved a ML of 4 and above on all topics is answering questions at the same difficulty level as those students who achieve the passing standard on the NCLEX exam.

**PassPoint**

PassPoint is a new product which incorporates the features of the PrepU NCLEX 10,000 with additional ways for students to practice and master content as they prepare for the NCLEX. Practicing in an authentic environment is an effective way to learn and PassPoint provides students the opportunity to take practice quizzes, as well as longer, simulated NCLEX exams.

Instructors can use settings to create and deliver a comprehensive NCLEX assignment of 75 questions or more across all NCLEX Client Need categories. This assignment will mimic the NCLEX experience for students and help gauge their current preparedness for the actual NCLEX test. The simulated NCLEX exams in PassPoint follow the NCLEX test design blueprint are designed to resemble the NCLEX in every way.

Previously, students were limited and could only take quizzes with a maximum of 25 questions. Now students can choose from a range of exams with the minimum number of questions (75) and the maximum number (265) aligned to the range they could see on the NCLEX. It’s important for students to have practice sitting in one place and answering a lot of questions and PassPoint gives students that experience. Students don’t know how many questions they will get when they sit for the NCLEX, so it’s important that the get some practice.

Pass Point also includes remediation links. These links take students to content related to their individual strengths and
weaknesses as indicated by their score on the simulated NCLEX or their day-to-day quizzing. The remediation links provide students chance to review content (but different content than they have seen up to that point) before taking more quizzes or practice exams.

As PassPoint becomes available in the fall of 2013 we will continue to gather both PassPoint usage and mastery data as well as student performance data to help us better understand the relationships between the two.

---


NCLEX Passing Standard Information from National Council of State Boards of Nursing.
