

NCBC Nurse Navigator Examination Standard Setting Report

Center for Educational Measurement, Excelsior College
December 2017

INTRODUCTION

The goal of this report is to outline the main activities that constituted the standard setting for the Nurse Navigators Exam. Standard setting, a critical and essential component of an assessment program, is a process by which an acceptable level of performance is defined (Kane, 1994, 1998). The most critical part of this process is to identify a cut-score which represents the standard of performance. In the case of the NCBC Nurse Navigator exam, the standard incorporated in qualitative statements of skills and knowledge that a candidate needs to function safely and effectively, as a Certified Navigator – Breast Nurse (referred to as Nurse Navigator in this report).

The Nurse Navigators exam is a paper-and-pencil that measures basic knowledge and skills expected of an individual seeking to become a certified Nurse Navigator. It is composed of 150 items, mostly multiple-choice item. A minority of the items were of the matching and the true-false types. Based on item analysis results, 120 items had been designated as scorable based on their difficulty level (p-value) and discrimination (point-biserial coefficient). The reliability of the exam as determined by the KR-20 coefficient was .86, with a standard error of measurement (SEM) 3.40. It is worth noting that the average item difficulty (p-value) was .86, and the lowest p-value was .45; the items on the exam were generally very easy for the examinees in the sample.

An important aspect of any standard setting is the validity of the process and its activities. The validity of the process cannot be ignored because doing so may raise issues of credibility and defensibility of the method used (Barman, 2008; Ricker, 2006). To this end, a series of virtual standard setting sessions were convened on November 16, 21, and 28, 2017. These sessions were used to translate a predetermined standard into a cut-score that represented the standard. The sessions were coordinated by NCBC and facilitated by the Center for Educational Measurement at Excelsior College.

Several cut-score setting methods are currently available for written tests. They are either norm-referenced or criterion-referenced. A norm-referenced method is used when only a fraction of candidates is needed to pass (Turnbull, 2009). As such, qualified candidates may fail or unqualified candidates may pass. A criterion-referenced (absolute) method such as those based on the idea of Angoff (1971) is based on the content of the exam and is independent of the performance of a cohort. Candidates may pass or fail depending on their knowledge and skills compared with clearly defined criteria (Turnbull, 2009). As a result, all candidates may pass or all may fail.

A modified Angoff method, a variant of the original Angoff (1971) method, was used for this standard setting. This criterion-referenced method is widely used in certification and licensure exams, and is well supported by research evidence (See, e.g., Plake, 1998, DeMauro and Powers, 1993). The modified Angoff procedure requires a panel of subject matter experts (acting as judges) to collectively agree on the definition of a *minimally qualified candidate (MQC)*, and then estimate the proportion of the MQCs who would answer a test item correctly.

This standard setting was conducted in accordance with guiding principles adopted by the *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council of Measurement in Education, 2014) with respect to 1) selection and training a panel of judges, 2) selection and execution of the standard-setting method, 3) provision of feedback to the panel, and 4) documentation of the findings. The sections that follow describe activities before and during the standard setting sessions.

PROCEDURES

In the Modified Angoff method, panelists are required to develop a common understanding of what the MQC should know and not know with respect to the domain being tested. To assist panelists in this task, a basic definition had been developed by the Nurse Navigators exam development committee ahead of time and offered to the panelists as a starting point. After much discussion, the panelists agreed on some changes and enhancements by suggesting attributes that they felt were reflective of minimally qualified candidate abilities.

Details of steps followed in conducting the standard setting are presented in Appendix 1. Briefly, panelist did a self-study on a standard setting PowerPoint training presentation and familiarized themselves with guidelines for defining the minimally qualified candidates ahead of three virtual meetings. At the first meeting, panelists were trained on the Modified Angoff standard setting process using practice items. After the first virtual meeting, the judges conducted round 1 ratings independently. The psychometrician then evaluated item rating discrepancy using the range and standard deviation. In the second virtual meeting, they discussed some items with the largest discrepancies. They were then instructed on how to use empirical p value to evaluate their own rating. Between the second and third virtual meetings the judges rerated all the items, this time using the p-values. At the third virtual meeting, again items with large discrepancies were discussed and judges had the opportunity to change their rating if they chose to do so.

Panelists completed a short oral survey during the first virtual meeting (to ensure that they understood their task) and a final survey after the last virtual meeting (to evaluate the overall standard setting process).