Standard Setting Report Summary

Certified Surgical First Assistant

November 2019

Submitted to:

NATIONAL BOARD OF
SURGICAL TECHNOLOGY
AND SURGICAL ASSISTING

Where people meet potential

Contents

Introduction

This report describes the standard setting study conducted for the National Board of Surgical Technology and Surgical Assisting (NBTSA) Certified Surgical First Assistant (CSFA) Examination. The study was conducted using methods that objectively facilitated the criterion-referenced nature of credentialing decisions, based on a method described by Angoff (1971)[1]. The study results were used to guide selection of a passing standard.

As with any professional practice standard setting process, some type of judgment is required. However, it is essential that the judgments involved in determining the passing point be made by qualified experts who are well informed regarding the intended use of the examination and possess the requisite knowledge and experience in the content domain to know what level of competence should be reasonably expected. Additionally, the judgments should be rendered in a meaningful way that accounts for the format and purpose of the test.

This determination must be made with recognition of the effects of potential error on classification decisions and the negative consequences of possible misclassification for examinees and the public. This report documents the appropriateness of the established passing point (cut score) used to determine examinee decision outcomes.

The underlying philosophy of the Angoff procedure is that the standard set relates to expectations of performance for those who are minimally competent (i.e., those with the requisite capabilities commensurate with the eligibility requirements). Therefore, this procedure requires judges to render an expected performance rating for each test item that reflects their expectation of performance for those who are minimally competent.

---

[1] Angoff, W.H. (1971). Scales, Norms, and Equivalent Scores. In R.L. Thorndike (ed.), Educational Measurement. Washington, D.C., American Council on Education, 1971.

Methodology

The judges serving on the standard setting study panel were selected by NBSTSA, all of whom were deemed to possess subject matter expertise. They were selected to provide for an appropriate balance on potentially relevant professional characteristics, such as area of special expertise, practice setting, and geographical distribution. See Appendix A for a listing of judges and their qualifications.

During a meeting in Olathe, KS on November 9, 2019, the judges participated in a standard setting study that consists of the following three major steps:

1. Definition of Minimum Competence
2. Rating of Examination Items
3. Consideration of Empirical Data

Judges engaged in an introductory presentation on the process that describes the standard setting activity and explains their role in the process. See Appendix B to see the presentation used.

Definition of Minimum Competence

In preparation of the rating process, a discussion regarding the definition of a minimally competent practitioner (MCP) was facilitated. An MCP is described as an individual who has enough knowledge to practice safely and competently but does not demonstrate the knowledge level to be considered an expert.

They then participated in a discussion regarding the definition of a minimally competent practitioner (MCP). An MCP is described as an individual who has enough knowledge to practice safely and competently but does not demonstrate the knowledge level to be considered an expert. The first step in this process was to identify the attributes of the MCP. A general discussion of these attributes occurred, and when the judges decided on a necessary attribute of the MCP, it was placed on a list. After considerable discussion, a suitable list of attributes was decided upon. The final list is presented in Appendix C.

Rating of Examination Items

Judges were then trained on the rating process. Central to this process is the notion that each rating is provided individually by each rater and reflect the answer to this question: What percentage of MCPs do you expect will answer this item correctly?

During this meeting the judges provided an Angoff rating, which is the percentage of minimally competent practitioners expected to respond correctly to an item. A rating was provided for each item on the test. Judges were advised that items selected for an examination very rarely perform below chance level (25% correct or less for 4-option multiple-choice items) or such that virtually all examinees answer correctly (95% correct or more).

Five items were used to orient the panel to the task before engaging in the rating of the remaining items on the examination form. The judges were instructed to provide a rating for each item (round 1 rating) after reading the stem and the response options.

Consideration of Empirical Data

Following the initial rating for each item, the answer key and p-value were presented so that the judges could re-evaluate their thought process and revise their ratings (round 2 rating). For example, the judges were specifically advised to consider the possibility that their ratings might be too high on items that they answered incorrectly when the initial ratings were recorded or if their expectations of performance for MCPs were significantly different from the p-value, which represents the performance of a sample of examinees who represent all levels of competency and performance. Judges then provide their round 2 ratings.

Ratings were then collected from the raters, and items for which the average rating was higher than the p-value by .20 points or more were identified as well as those in which the highest and lowest ratings differed by .20 points or more. The judges discussed these items to determine why their expectations differed significantly from the difficulty for all examinees or from each other. All raters were then given the opportunity to revise their ratings for the subset of items identified. There were no changes in the ratings resulting from this discussion.

Results

Summary of Ratings

The table below presents the results of the analysis of the judges' ratings. The judges' individual cut score estimates (round 2 ratings) ranged from 61.44% to 68.12%. The mean of the judges' estimates yields a cut point of 64.94% (97 of 150 items). The standard error of the judges' ratings was calculated to be .61. The standard error of the ratings represents the expected amount of variability in ratings if the judges were to repeat the process and can be interpreted as a confidence interval around the judges' final estimate. The individual judges' ratings can be found in Appendix D.

JUDGE #	MEAN RATINGS ROUND 1	SD OF RATINGS ROUND 1	MEAN RATINGS ROUND 2	SD OF RATINGS ROUND 2
1	65.10	12.87	62.83%	14.05
2	61.44	13.61	61.93%	13.75
3	67.01	11.07	66.20%	11.61
4	66.58	14.07	66.50%	13.78
5	66.28	12.28	65.93%	12.47
6	68.12	16.96	67.70%	17.01
7	67.08	14.12	66.50%	14.43
8	65.17	12.85	64.73%	13.68
9	63.46	13.57	62.70%	14.43
10	65.10	14.66	64.33%	14.70

Established Passing Standard

The results of the study were presented to NBSTSA on November 9, 2019. While the mean of the ratings was selected, the NBSTSA was also advised that it would be acceptable to establish a cut score within one of the provided confidence intervals. Following a discussion, NBSTA established 97 out of 150 total scored items as the raw score passing point for the CSFA examination. The selected passing point of 64.94% (97.41) was the mean of ratings, and a rounded raw cut of 97 was therefore selected.

The form used for standard setting will be released on January 2, 2020 and will serve as the base form for the development of subsequent forms of the examination, to be made parallel and equivalent using statistical pre-administration equating and adherence to the content allocation requirements of the examination specifications.