PROMETRIC

Standard Setting / Cut Score Study Report

Test Development
Solutions

Child Life Certification Commission (CLCC)
Certified Child Life Specialist (CCLS)

Standard Setting / Cut Score Study Procedures & Results

Date:
May 3, 2024

Submitted by:
Nick Baker
Assessment Design Specialist

Prometric
1501 South Clinton Street
Baltimore, MD 21224

INTRODUCTION

This report documents the procedures employed to select the standard setting (cut score) study panel, the methods used in conducting the study, and the analyses performed for the standard setting study conducted remotely for the Certified Child Life Specialist (CCLS) examination on April 23 and May 1, 2024.

STANDARD SETTING METHODOLOGY

There are several recognized processes for standard setting. For the CCLS standard setting study, Prometric recommended and adopted the modified Angoff method. The modified Angoff method requires a minimum of eight subject matter experts (SMEs) to reach consensus on a definition of the minimally competent candidate. SMEs then read and answer each item in the test and predict the proportion of minimally competent candidates who would answer each item correctly. Resulting ratings are analyzed and discussed with the participating SMEs. The SMEs are given an opportunity to affirm or revise their original ratings. Summing the ratings across items and across SMEs provides a recommended cut score. Prometric facilitated a two-day remote meeting on April 23 and May 23, 2024. The results of this process are provided to CLCC within this report, with the final cut score to be determined by CLCC.

SELECTION OF THE PANEL OF JUDGES

CHARACTERISTICS

The Child Life Certification Commission (CLCC) recruited a panel of judges from CCLS professionals. A panel of twelve judges completed the standard setting study. The agenda for the study can be found in Appendix A.

REPRESENTATIVENESS

All twelve of the judges on the panel held the CCLS credential. Eight, of the judges held Master's degrees and four held Bachelor degrees. The panelists represented six U.S. States, two Canadian Provinces, and one Zambian Province. The panel reported an average of 8.66 years of experience (SD = 5.45). More information on the composition of the panel can be found below in Table 1.

CONDUCT OF THE MEETING

Prior to the meeting, panel members were provided with a document describing the purpose and procedures of the cut score study and an agenda for the meeting.

At the beginning of the first meeting, Prometric facilitated a general orientation session with the panel members, which included discussion about the need for a standard setting exercise, and the steps the panel would be taking in conducting the study. Questions and comments on the procedure were entertained with emphasis on the importance of group discussion and participation.

DEVELOPMENT OF THE STANDARD OF MINIMUM COMPETENCE

Following the description of the procedure, the panel developed a standard of minimal competence identifying what a minimally competent certified child life specialist would know and what they would find challenging. The test specifications were presented to the panel as a starting point for the discussion and articulation of the attributes of minimal competence. This panel systematically went through the test specifications and identified concepts/content that the minimally competent CCLS would know and would find challenging. These discussions resulted in a final standard of minimal knowledge, skills, and abilities (KSAs) composed of specific statements about what the candidate needed to know to be considered minimally competent for certification.

PRACTICE RATINGS

After the panel agreed on the standard of minimal competence, they then completed practice ratings of a sample of multiple-choice test items (selected from the actual examination) following a modified Angoff procedure. During the ratings, Prometric repeatedly reminded the panel members that the purpose of rating the test questions was to judge how many out of 100 minimally competent candidates would answer each question correctly. They were not to judge the average (above minimally competent) candidate, nor were they to judge the amount of knowledge experts would possess. These points were repeatedly emphasized in the general discussion session as well as in panel rating sessions to avoid rating errors. No key was provided, and the panel members were asked to answer each item and rate "How many out of 100 minimally competent candidates do you think WILL answer this question correctly?"

Following the rating of all sample questions, each participant's rating was aggregated and entered into a spreadsheet feedback form and discussed. Panel members were directed to give a second rating following the discussion if they were inclined to change the first rating based upon their discussion of what the minimally competent candidate would know.

RATING OF THE ITEMS

Following the first meeting, the panel members were asked to take and rate (or judge) the items on the 125-item multiple-choice test. The panel was instructed to read each question, answer the question the best they could, and rate each item as to "How many of 100 minimally competent candidates would get this item correct?" No key was provided and there was no discussion among committee members.

After completing this task, each panel member's ratings were entered into a spreadsheet and their individual recommendations for the cut score were computed. The group's average cut score recommendation was also computed as were the confidence interval adjustments associated with the panel recommended cut. Each panel member's test answer sheet was scored according to the key developed by the CLCC and Prometric.

Once these processes were completed each panel member received their scored answer sheet with the total obtained score and individual cut score recommendation, and rating sheets. As a 'reality check' following the first rating the judges were asked to look at their recommended cut scores, and their obtained scores on the test, to see if they would have failed themselves on the examination following their first rating. A spreadsheet showing each panel member's obtained score on the examination and their first individual recommended cut score was presented to the panel (only the obtained score and cut score were shown; all identifying features of the panel members were removed and the sequence altered to protect the anonymity of the panel members). The mean obtained scores of the panel members and the current panel recommended cut score were also presented on the overhead screen. The members were encouraged to compare their first individual recommended cut scores to their obtained scores and to the obtained and recommended scores of their fellow panel members. Any members who would have failed themselves were encouraged to give special consideration to the discrepancy between their obtained scores and recommended cut scores.

The feedback report containing the panel ratings on each of the 125 items, stored on the computer, was viewed on a screen for panel discussion of first ratings. For each item the feedback report contained the key and the frequency of ratings from the panel in categorical intervals of 20 percentage points. In addition, for approximately every 5th item the percentage correct statistic was provided.

The ratings for each item were examined in sequence and the panel was encouraged to discuss their differences in ratings, particularly when there was large variation in the panel judgments. During this discussion panel members were asked to reexamine their original ratings in light of the discussion and data. They were instructed to take into consideration (1) the anonymous judgments of the entire group; (2) the standard of minimal competence; (3) whether they answered the item correctly; (4) the comments of the group, and (5) the percentage correct statistic when provided. Then the panel was given an opportunity to change their rating, if desired, to better reflect the group definition of what the minimally competent candidate would know.

After updating the spreadsheet with the panel members' adjusted ratings, the group's average cut score recommendation was recomputed, along with the confidence interval adjustments associated with the panel recommended cut. The panel was presented with the final panel recommended cut score and confidence interval ranges .