CRC® Examination
Pass Point Setting Procedure

The International Foundation for Retirement Education (InFRE) commissioned a Pass Point Study for an examination form of the Certified Retirement Counselor® (CRC®) examination program in order to determine a fair and defensible standard for the exam. Personal Touch Assessment (PTAssess) facilitated the Pass Point Study which was held virtually on March 30, 2016 with nine subject matter experts. Best practice indicates a standard setting study be conducted whenever a practice analysis and new examination blueprint become effective.

A. Guiding Principles

The design and execution of the Pass Point Study for the CRC® Examination was consistent with procedures adopted in the *Standards for Educational and Psychological Testing* (AERA, APA, NCME, 1999) and with standards published by the National Commission for Certifying Agencies (NCCA). Relevant practices recommended in these documents were applied to study activities relating to the selection and training of judges, selection and implementation of the standard setting methods, provision of feedback to judges (through practice rating activities), and documentation of the findings.

B. Selection of the Standard Setting Method

Standard setting activities in most credentialing examination programs are governed by two general approaches: norm-referenced procedures, which yield a **relative** test standard and criterion-referenced procedures, which yield an **absolute** test standard.

The norm-referenced approach establishes the passing point based upon candidate test scores and sets the percentage of passing candidates, thus establishing a "passing rate" of say, 75% as an example. The "relative" nature of this type of standard setting means that the actual passing point (i.e. the number of questions one must answer correctly in order to pass) may shift in relation to the proficiency level of the candidate cohort sitting for the test, while the percentage of successful candidates remains constant.

In contrast, criterion-referenced procedures determine the passing point based on a predetermined or "absolute" level of mastery required to achieve a passing score. Thus, the level of proficiency associated with the passing point remains constant across different forms, while the passing rate varies in relation to the ability level of the candidate cohort.

Criterion-referenced methods are favored in certification and licensure testing. Of the available criterion-referenced standard setting techniques, the procedure attributed to William Angoff (1971) is the most widely used in the credentialing industry. Angoff suggested that expert judges review each item on a test, and then make an estimate of the probability that a "borderline" passing candidate would answer the item correctly. The average of these

probabilities, aggregated across judges, would be the minimum passing score on the test. The Angoff technique has been modified since its inception to improve the technical quality of the resulting standard.

Best practice dictates that multiple rating iterations be conducted in standard setting studies to arrive at the most reliable results. The process described below used two iterations of rating exercises.

C. Implementation of the Certification Standard Setting Method

A step-by-step description of the procedures used at the standard setting meeting is presented below.

Step 1. Judges were provided with information about the exam (for example, exam specifications), the standard setting process and the concept of the "barely qualified" candidate in advance of the meeting. Judges also reviewed the content of the examination forms. This review allowed judges to become familiar with item difficulty, the range of test content, types of test items, and the correct or best answer to each item.

Step 2. PTAssess started the session(s) by providing an overview of the Pass Point Study procedure and process.

Step 3. Judges discussed and defined the profile for the "minimally competent" or "barely qualified" practitioner who has "just enough" knowledge and skills to merit certification. This discussion was based on definitions of several responsibilities from the domains in the CRC® Content Outline. The purpose of this activity was to develop a consensus among judges of the proficiency level of individuals who would represent the "minimally competent" candidates. During the subsequent rating activities, judges relied on these consensus profiles to make item performance estimates.

The discussion began with a summary of the eligibility requirements for the certification exams. The judges agreed that the minimally competent candidate would have undertaken reasonable exam preparation.

Step 4. The Angoff (1971) methodology, which is the most widely used criterion-referenced passing point technique within the credentialing industry (*Sireci & Biskin, 1993; Fidler,1996*), was then introduced. It is based on the judgments of content experts regarding the expected test performance of candidates who are minimally competent.

Step 5. Judges applied the Modified Angoff technique to a sample of items from the CRC® examination representing each domain. The objective of this activity was to ensure that the judges fully understood the rating task and the consequences of their item estimates.

Step 6. Judges applied the rating technique to the 175 items which comprise the scored items on the CRC® exam form. Judges had the benefit of item statistics and worked independently to avoid rater bias.

Step 7. Judges were directed to provide a second set of ratings. These were Global Estimation Ratings based on the test as a whole. These are designed to establish acceptable range values for the passing points and resulting passing rates derived from the Angoff rating session.

D. Data Analysis Activities

Modified Angoff. The Modified Angoff scale required judges to select a percentage from a series of incremental value ranges: 10% – 19%, 20%– 29%, 30% – 39%, and so forth, through a final scale point of 100%.

The procedure for computing estimated passing scores for the Modified Angoff rating session was as follows. Each individual judge's item ratings were first summed across items, and then a grand total was computed by summing across judges. Finally, the grand total was divided by the total number of judges in the group. To calculate passing points, the average percentages were multiplied by the number of items in the test.

The Judges make a recommendation, and InFRE makes the final decision. Determining a pass point for an examination is not just a result of an analysis of data and acceptance of Judges' recommendations. Historical trends, policy concerns, and possible rating biases on the part of Judges must be factored in as well.

In conclusion, the psychometric procedures for standard setting using a legally defensible method to set the CRC® passing point standard were carried out with fidelity and any adjustments to the standard were based on policy consideration and the concern about perceived relative cost regarding false positives and/or false negative decisions (not awarding a credential to someone who is qualified because the score is too high and/or awarding a credential to someone not qualified because the score is too low).

References

AERA, APA, and NCME. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational Measurement* (2nd ed.) (pp. 508 – 600). Washington, D.C.: American Council on Education.

Fidler, J. R. (1996). A survey of examination-related practices employed by credentialing agencies. Unpublished manuscript.

Sireci, S. G., & Biskin, B. H. (1993). Measurement practices in national licensing examination programs: A survey. *CLEAR Exam Review*, 21-25.