This document identifies the proposed revisions to the NCCA Standards as a result of the 30-day public comment period that occurred from July 1 – July 31, 2021. This document is not a copy of the complete Standards, rather, only includes essential elements and commentary with revisions as a result of the noted 30-day public comment period. Only the marked changes in tracked changes are open for review and comment during the current 30-day public comment period, August 31 – September 30, 2021.

Contents

Standard 1: Purpose

The purpose of the certification program must be to recognize each individual who meets established criteria. These criteria must uphold standards for practice in a profession, occupation, role, or specialty area.

Commentary:

1. Certification can be offered for a specific profession, occupation, role, or specialty area across multiple disciplines. The program should specify the target population(s) for certification, including the level of experience for the targeted practitioner, as well as the purpose of the certification program.

2. For purposes of the *NCCA Standards*, a **credential** is defined as a "formal recognition awarded to an individual who has met predetermined standards and maintains any renewal requirements" and **designation** is defined as: "An indication of a credential that an individual holds, which could be a specific title, letters, or acronyms before or after an individual's name."

3. ~~In most cases, the designation issued or formally recognized by the certification program aligns with the credential. In rare cases, however, the program may issue a designation that does not accurately reflect the credential. In those cases, the certification program should:~~
    *   ~~provide a rationale with supporting information that addresses the potential misunderstanding and/or misuse of the designation by stakeholder groups such as employers or consumers of a certificant's services, and~~
    *   ~~have a mechanism to identify variations among certificants that is publicly available.~~

4. ~~4.~~3. In some cases, a certification program may not offer a designation, meaning there is no formal indication of the credential such as a specific title, letters, or acronym that certificants can use before or after their name. In such cases, the certification program should provide an explanation.

5. ~~5.~~4. Suggested evidence to document that this Standard has been met may include a mission statement, bylaws, candidate handbook, policy and procedures document, and other publicly available documents.

    *For the purposes of *NCCA Standards*, wherever the term "publicly available" appears, it should be interpreted as "available without request." If any information identified as "publicly available" in these Standards is only available upon request, the certification program should provide a rationale.

Standard 7: Program Policies

The certification program must establish, enforce, and periodically review certification policies and procedures related to certification and challenges to certification decisions.

Commentary:

7. Procedures for requesting accommodations should be stated clearly and be publicly available.

   The procedures should include mechanisms that will ensure that proper evidence is submitted to the program to assist it in making a determination regarding the requested accommodation.

   Any accommodation provided should be reasonable and not compromise the fundamental nature of assessment or the validity of the certification decision. Certification programs should not reveal on score reports or certificates that any accommodation was provided during the administration of the examination.

   ~~Examples of~~Certification organizations should comply with applicable laws and regulations. Examples include the Americans with Disabilities Act for organizations operating in the United States and American entities operating outside of the United States (with respect to accommodation requests, accessibility of digital content and exam administrations, and confidentiality of medical information), nondiscrimination laws, antitrust laws, applicable laws that govern the industry or profession, and other relevant provisions.

Standard 13: Panel Composition

The certification program must use panels of qualified subject-matter experts (SMEs) to participate in job analysis, item development, standard setting, scoring, and other examination-related activities.

Essential Elements:

A. The certification program must document the responsibilities entrusted to each panel.

B. The process of recruitment and involvement of panelists must be documented and must prevent the undue or disproportionate influence of any individual or group.

C. The certification program must define the relevant demographic and professional characteristics of panelists and the target composition of each panel. Each panel must represent the relevant characteristics of the population to be certified as the program defines them.

D. The certification program must document the characteristics of the panelists for each panel.

E. Documentation of meetings and activities must include decisions and recommendations of panelists.

Commentary:

1. For the purposes of *NCCA Standards*, a panel refers to people working collectively or individually on tasks for a certification program; panels may be known by other names, such as committee, team, council, commission, or group. Generally, panels should represent the population being certified, and selection should take into account diversity, equity, and inclusion whenever possible. Most SMEs should be certified in the discipline and/or actively practicing; however, individuals who are qualified in other disciplines may serve as panelists. SMEs' levels of experience and knowledge should be congruent with the activity in which they are engaged and may therefore include some newly certified individuals. Panelists may also include supervisors, faculty, and regulators.

2. The members of each panel should be provided with information regarding the purpose of the examination, the role and expectations of the panel, the rules governing panelists' participation, and a general description of the activities in which they will be involved.

3. Individuals may serve on more than one panel, and they may serve for several years; however, certification programs should ensure that there is fair opportunity for a broad range of SMEs to participate over time. A system of terms of service that includes a rotation schedule for panel membership is a useful means of ensuring broad input into the examination program. *(Note – this last sentence in green was not revised. Rather, it was moved from commentary 1 and incorporated as a part of what is now commentary 3. This is not an item open for feedback during this public comment period.)*

4. Suggested evidence to document that this Standard has been met may include the following: procedures and requirements for the selection of qualified individuals for the panels; lists of panelists along with their key characteristics related to the purpose of the panel on which they are serving, and panel meeting minutes.

Standard 19: Scoring and Score Reporting

The certification program must employ and document sound psychometric procedures for scoring, interpreting, and reporting examination results.

Commentary:

1. Certification programs should establish and describe quality-control procedures for monitoring the accuracy of calculations used to produce scores and the conversion of raw scores to standardized, equated, or scaled scores. The organization should clearly document the weighting of items or tasks. The scale on which scores are reported should support interpretations that are consistent with the purpose of the examination.

2. For performance and other examinations where responses are scored by judgment, developers should document methods for developing scoring rubrics, judging responses, reducing rater bias, and increasing inter-rater agreement and consistency to ensure an acceptable level of consistency in scoring judgment-based items. If the performance component is scored without raters (e.g., computer scored), the program should demonstrate how successful performance is being evaluated accurately. Types of documentation to support these items may include the following:

   - criteria used for selecting judges;

   - a description of the materials and methods for training judges;

   - evidence demonstrating that the primary source of variation in candidates' scores comes from their performances, not rater or computer scoring error; and

   - summaries and results of process, rater, or score audits or other technical controls to ensure that the candidates' performances are the primary determinant of whether they pass or fail examinations.

3. The certification program should provide candidates with an explanation of the types of scores reported, appropriate uses, and potential misuses of reported score information. Information about performance in relation to the passing standard provided to failing candidates may be quantitative or qualitative and should help candidates make informed decisions related toabout retesting. Feedback should be appropriate for the type of examination.

4. If domain-level information has low reliability, programs are advised against reporting it to candidates and other stakeholders. When domain-level or other specific feedback is given to candidates, the certification program should provide estimates of its precision and/or other guidance.

5. The certification program should ensure the fairness of the examination for all populations. If the program detects potential unfairness, it should take steps to understand its causes and, if possible, remedy it.

6. Suggested evidence to document that the Standard has been met may include descriptions of scoring procedures, training documents, and quality-control procedures, such as the following:

   - security procedures pertaining to scoring, reporting scores, and maintaining score records;

   - quality-control procedures pertaining to scoring, reporting scores, and maintaining score records (checklists, policies, narrative);

   - sample score reports for passing and failing candidates, including instructions on interpreting feedback that is provided;

- policies, procedures, and supporting materials for scoring objectively scored examinations; and

- policies, procedures, and supporting materials for judgment-based scoring (e.g., procedures, required number of judges, development of and training on scoring rubrics).

Standard 20: Evaluation of Items and Examinations

The certification program must evaluate items and examination forms to ensure that scores are sufficiently reliable for the decisions that are intended.

Essential Elements:

A. Certification programs must evaluate item performances as well as calculate and report estimates of score reliability, decision consistency, and standard errors of measurement using methods that are appropriate for the characteristics of the examination.

B. Estimates of score reliability and decision consistency must be reasonable to support accurate pass/fail decisions. If the certification program makes pass/fail decisions based on subscores (i.e., the assessment is multiple-hurdle, or non-compensatory), the reliability of each subscore for which a pass/fail decision is rendered must be reasonable.

C. When examinations are translated or otherwise adapted for use by subgroups of the target population, across languages and/or cultures, certification programs must demonstrate that results obtained from adapted and source versions are comparable.

D. For performance examinations, certification programs must demonstrate that results are equivalent across raters and performance tasks.

Commentary:

1. The selection of reliability statistics required for an examination depends on the type of assessment and the purpose of the scores. Programs should document the reliability estimate(s) and provide a rationale for the methods used. Examples of such methods could include inter-rater agreement, inter-rater consistency, agreement between computer scoring and raters for performance examinations, and/or internal consistency estimates.

2. If a program makes decisions using domain-level information, it should demonstrate that the reliability of that information is sufficient and provide a rationale for how it weights and uses domain-level information.

3. When candidate volumes are so small or there are other factors which lead to reliability estimates that are not meaningful, programs should describe the procedures used to demonstrate that the decisions made on the basis of scores are reasonable and fair.

4. There should be evidence that translated or adapted examinations are testing the same construct as in the original examination. Simple translation and back-translation are not ~~adequate~~sufficient. ~~When candidate volume is sufficient to permit the analysis, differential item functioning (DIF) studies should be used to demonstrate that the construct is equivalent across the two versions. A DIF study indicates the extent to which examinee subgroups have systematically different correct response frequencies while controlling for their ability levels.~~ Additional evidence regarding the processes (e.g., reviews by bilingual SMEs, dual language display of items) and/or statistical analyses (e.g., differential item function (DIF) studies, factor analysis, descriptive summary statistics) should be provided.

5. Examination evaluation information should include such things as item analysis, reliability, decision consistency, speededness, item response time, and candidate feedback. This evaluation should be conducted frequently enough to ensure integrity of examination results.

6. Suggested evidence to document that this Standard has been met may include the following:

    -   reliability coefficients, overall standard error of measurement, information function, and/or other statistics pertaining to the consistency of scores;

    -   indices of classification consistency, conditional standard errors of measurement, or other measures of score consistency at the cut score; and

    -   information about how non-compensatory domain-level scores and other measures are evaluated and combined.