

08.10.2017

Technical Memo

To
Indiana State Board of
Education

From
**Indiana Technical
Advisory Committee**

Karla L. Egan, TAC
Chair; Derek Briggs,
Chad Buckendahl,
Gregory Cizek,
Marianne Perie,
Ed Roeber

CC
Cynthia Roach,
Charity Flores

Re
CORE Assessments

This memorandum responds to a request from the Indiana State Board of Education to provide the Evaluation System Group of Pearson (Pearson) with additional guidance regarding the psychometric characteristics of the CORE Assessments. The IN TAC was asked by the IN Department of Education (IDOE) to review the technical characteristics of the Core Assessment, Indiana's educator licensure examinations. Pearson first presented information on the Core Assessments during the April 2017 meeting of the Technical Advisory Committee (TAC). In response to requests for additional information by the TAC, Pearson again presented information regarding four sources of evidence about the Indiana CORE Assessments. These sources of evidence focused on content validation, strategies for responding to low volume fields, reliability, and reviewing passing standards. It was the opinion of the TAC that the information presented by Pearson was not adequate for making recommendations to or decisions by the IDOE. Thus, the purpose of this memo is to provide more explicit recommendations regarding acceptable validity evidence in support of the CORE Assessments. The TAC would be pleased to respond to any questions related to this memo before work is begun or presentations planned for a future TAC meeting.

As background, the steps for developing credentialing (e.g., licensure, certification) examinations such as the CORE Assessments involve defining the purpose, job (practice) analysis, content review, bias review, field testing, psychometric analysis, forms assembly, standard setting, and program-specific maintenance activities. In the following sections of this technical memo, we describe the key sources of evidence that we believe would be needed to provide support for the intended interpretations and uses of scores from the CORE Assessments. We have focused this memo on three sources of evidence: content validation, reliability, and standard setting.

To avoid redundancy, we note here that, for each of these sources of evidence, appropriate documentation would include: information about the qualifications of the subject matter experts (SMEs) who participated in any procedures; information about the processes and procedures that were implemented; the results of those processes; and any decision rules or revisions that were made based on the results. An additional consideration for the CORE Assessments is the methods that would be appropriate are related to the volume of test takers in a given field. That is, some fields have larger volumes of candidates than others, and this consideration affects the validation strategies that would be appropriate.

RECOMMENDATION 1: Contractor plans for addressing sources of validity evidence should reflect strategies that are appropriately aligned with the volume of test takers in a respective program.

Content Validation

For credentialing examinations like the CORE Assessments, content validation is the primary source of evidence that connects the assessment to expectations for practice. Procedures called *job analyses* (or *practice analyses*) are considered necessary for legal defensibility and are uniformly regarded as best practice within the testing industry. A job analysis is a systematic process that is used to define the knowledge, skills, and abilities judged necessary for competent entry-level practice in a profession. Although there are different methodologies that can be applied given the domain and the size of the target population, a common methodology is to develop a series of statements (e.g., tasks, competencies) that represent important aspects of the

domain. These statements are evaluated by a committee of subject matter experts (SMEs) and then formatted as a questionnaire that is then used to survey practitioners to determine the extent to which the statements reflect current, emerging, and declining practice in the field.

RECOMMENDATION 2: Documentation of the draft task statements or competencies; the tasks or competencies that characterize a given domain; the sampling plan for a survey of practitioners for each domain; and the data analyses, results, decision rules, and how the results were translated into the test blueprint that are used to develop CORE Assessment forms should be provided by the contractor. Such documentation would, necessarily, be provided separately for each field for which a CORE Assessment is administered.

If a survey of practitioners is not used (e.g., in the case of a low volume field), documentation of the rationale for the job analysis design and implementation would be required as part of the technical documentation for the program. Evidence for each element in a common job analysis process is briefly noted here:

Review of draft task statements or competencies – Technical documentation for this activity would include evidence of the review criteria in the context of the target population (e.g., importance, accuracy, job-relatedness, entry-level). Documentation would further include evidence of the development process, recruitment and qualifications (e.g., representativeness) of the participants who reviewed draft versions, process, and results of the review with revisions suggested by the panelist.

Questionnaire development – Evidence for this element would include copies of the questionnaire that was developed for the survey of practitioners. The questionnaire should align with the task statements or competencies that were previously reviewed. Information about the rating scale(s) that were used to collect respondent ratings would also be part of this documentation. It would also be appropriate to include a description of the survey methodology to obtain representative responses from the target population of practitioners.

Sampling design – Because each field has a different population, the sampling frame should be applied to the respective field and not based on a single process. This is particularly important given the wide range of volumes for the CORE Assessments. When appropriate given the size of the population, the purpose of a survey in the job analysis process is to collect a larger sample of representative feedback from practitioners beyond what a smaller focus group or committee can provide. When evaluating the sampling design and results, information is needed about the target population. Documentation for this part of the process would include information about the size of the target population along with information about the demographics (e.g., sex, race) to be able to later evaluate the representativeness of the response. Specifically, this information can then be used to determine the sampling error associated with the survey results or whether any post hoc weighting of the results based on population characteristics occurred or would be needed.

Analysis and translation to blueprint – For this part of the process, we would expect to see a description of the data analyses that were applied to the survey respondent information. In addition, the process and decision rules that were used to then translate the results of these analyses into the assessment

blueprint would be necessary to be able to link the evidence of job-related practice from practitioners to the assessment. When there are deviations from a direct link from the results to the blueprint, the technical documentation should include a justification or rationale (e.g., emerging practice, declining practice) to support the alternative weighting.

Because content validation evidence is of primary importance for a credentialing examination, the plan, implementation, and documentation for the job analysis is a critical component of a validity argument for the CORE Assessments.

RECOMMENDATION 3: In addition to the evidence from the job analysis, a detailed written summary of the Content Advisory Committee (CAC) and Bias Review Committee (BRC) plans and procedures should provide be provided to document the people, process, results, and decision rules used for each CORE Assessment field. (NB: A generic description for the collection of CORE Assessments is not acceptable.)

Reliability

An important source of evidence that lends greater or lesser confidence in assessment scores or decisions is the reliability or dependability of scores earned by candidates on the assessment. A goal of assessment is to eliminate systematic error (i.e., bias) and to minimize random error so that a reported score is a dependable representation of the candidate's ability. Different methodologies can be used to estimate these different sources of error. Reliability as estimated using Cronbach's alpha is often a key source of evidence; however, it may not always be the most important or relevant

statistic for credentialing examinations. INDOE should expect measures of internal consistency of 0.80 or better and measures of decision consistency of 0.85 or better. For forms that do not meet these expectations, explanations should be provided along with plans for improvement.

RECOMMENDATION 4: In addition to traditional internal consistency reliability, evidence of decision consistency in addition to estimates of scorer or rater error from the scoring process for constructed response questions should be provided for each CORE Assessment.

When reliability evidence is presented in technical reports or manuals for the program, additional narrative discussion is needed when the values do not support assertions of the reliability of the scores, scorers, or decisions. This discussion would include corrective action plans to improve the evidence or explanation for instances that can occur for very low volume fields about why target values may not be achieved.

Standard Setting

Standard setting is the process used to obtain cut scores on tests. In educator credentialing examinations such as the CORE Assessments, the procedures used and results of standard setting are a critical aspect of the credentialing process that must be documented.

RECOMMENDATION 5: A policy establishing periodic review of passing scores for examinations should be adopted for the CORE Assessments.

These policies will often define the criteria associated with when to revisit the passing scores and will frequently correspond with the development or redevelopment of a given field (e.g., every 5 to 7 years, when there are significant changes to the content, significant changes in the pool of

candidates, significant changes in curriculum or instructional practices at state educator preparation programs, or significant changes in credentialing policy in the state).

RECOMMENDATION 6: The cut scores used for CORE Assessments should be reviewed in combination with external validity evidence that may include performance disaggregated by institutions, course taking patterns of candidates, and candidate subgroups (e.g., traditional pathway, alternative certification).

These additional analyses may be useful to policymakers in their evaluation of the reasonableness of the passing standard when comparing expectations to observed performance.