AMERICAN INSTITUTE OF CONSTRUCTORS
CONSTRUCTOR CERTIFICATION COMMISSION DOCUMENT NO. 17

QUALITATIVE AND QUANTITATIVE EVIDENCE COLLECTED FOR VALIDATION OF THE CONSTRUCTOR CERTIFICATION PROCESS

What Qualitative and Quantitative Evidence has been collected for the Examination Process?

Downing and Haladyna (1997) in their article titled, Test Item Development: Validity Evidence From Quality Assurance Procedures published in the *Applied Measurement in Education* journal (p 61 - 82) assert that "the gathering of supporting evidence for validating a specific test use or interpretation must begin with a careful and systematic approach to the task of creating the test items." They suggest that all certification "programs can and should attend to these activities to the extent that their resources allow" (p 63). The eleven types of quality evidence they identified in the item development stage are stated below.

1. Item Content Definition.
2. Test Specifications.
3. Item-Writer Training.
4. Adherence to Item-Writing Principles.
5. Cognitive Behavior.
6. Item Content Verification.
7. Item Editing.
8. Bias-Sensitivity Review.
9. Item Tryout and Pretest.
10. Key Validation and Verification.
11. Test Security Plan.

In addition, Downing and Haladyna (1997) insist that this "ideal process" also contains a quantitative component which pertains to the desirable characteristics of the item responses. They suggest that there are several types of quantitative evidence that can be used to validate item responses using the item analysis report. The types of quantitative evidence conducted during the item analysis stage are identified below (p 73).

1. Key Validation by Subject Matter Experts.
2. Item Factor Analysis Connected to the Item Classification and Test Specifications.
3. Distractor Functioning and Plot the Item Response Patterns for each item.
4. Differential Item Functioning (DIF) for Deviations among Selected Groups.

The qualitative and quantitative criteria outlined by Downing and Haladyna (1997) above are consistent with the numerous standards for educational and psychological testing. The criteria outlined above are recommended according to numerous measurement associations such as American College Testing (1998); American Psychological Association (2002); American Educational Research Association (2001); Council on Licensure, Enforcement, And Regulation (1994); National Organization for Competency Assurance (1993); National Council on

Measurement in Education, (1999). Therefore, the Certified Professional Constructor in

conjunction with our testing consultants National Assessment Institute and Professional Testing Corporation have implemented these criteria or procedures to ensure that a valid, reliable, and defensible examination process exists. Attached at the end of this paper are two detailed tables. One titled Qualitative Evidence in the Item Development Stage and the other titled, Quantitative Evidence in the Item Analysis Stage. The qualitative table summarizes Downing and Haladyna's (1997) eleven types of evidence, the specific activities that are performed to gather the evidence, and the resulting data required to establish qualitative evidence for the item development stage. The quantitative table outlines the test item analysis conducted after each examination administration. Both tables are shown in Appendix A on page four.

Numerous experts (Downing, (1997); Haladyna (1987, 1994, 1997); LaDuca (1994)) contend that the primary objective of a high-stakes certification program is to establish various forms of objective evidence throughout the process to ensure that the examinations reflect the knowledge and skills necessary for competent performance in managing the construction process. One form of evidence is done by providing proof of the validity of the examination process by collecting evidence during the role delineation, the test specification, item development, item banking, pilot testing, test assembly, and test interpretation stages of the process. Haladyna (June 1987) in his article titled Three Components in the Establishment of a Certification Testing Program published in *Evaluation & the Health Professions* states that "other requirements for granting certification may include" collecting evidence of adequate formal training, practical experience, and continuing education following formal education (p 140).

Certification in construction is intended to recognize which candidates in the construction profession can perform at or above an entry level with respect to the knowledge and skills identified as necessary for competent practice. These high-stakes Certified Professional Constructor (CPC) Examinations have significance consequences for the candidates, the profession, and potentially the public. Due to the high stakes nature of these examinations, it is the responsibility of the AIC Constructor Certification Commission to be able to ensure that the examinations reflect the knowledge and skills necessary for competent performance. Hence, test validation assumes that evidence will be collected concerning the relationship between the test content and job performance. This is met by creating a link between the test items and the skills performed. This has been done through the development of the test specifications.

What is the role of the Constructor Certification Testing Agency?

Our testing agency provides numerous elements for a sound certification testing program such as validity, reliability, and defensibility. The primary objective for obtaining a professional testing agency was to ensure a valid, reliable, and defensible certification process and to utilize their expertise in gathering and collecting objective evidence for the Constructor certification process. Validity refers to the appropriateness, meaningfulness, and usefulness of specific inferences made from test scores according to the American Psychological Association, American Educational Research Association, and the National Council on Measurement in

Education (1985). These associations claim, "test" validation as the process of accumulating evidence to support such inferences. The inferences regarding specific uses of a test are validated, not the test itself" (p 9). Downing and Haladyna (1997) believe, "validity is the most important consideration in test evaluation" (p 61). They emphasize, "validity" of a candidate's examination score involves collecting and organizing evidence to support a specific test score interpretation," and "a primary type of validity evidence derives from the item development process and the item responses" (p 61). Reliability is the degree to which test scores are free from errors of measurement.

The legal basis for certification is an effort to maintain high standards for the profession and to maintain a high quality of service provided to the public by members of the profession. This implies that each certification testing program must provide a body of evidence to support the use of an examination as an inference regarding professionals made from test scores. Therefore, it is imperative in a certification process that a careful, well-planned, systematic item development and item analysis process be established to provide qualitative and quantitative evidence for awarding the certification. This is the primary purpose for hiring a professional testing agency not affiliated with any university.

A testing agency was hired to review the role delineation study, assemble a group of Subject Matter Experts with the sole purpose of developing the test specifications. They are also involved in conducting multiple-choice item writing sessions to train item writers. After the multiple-choice item writing sessions are held, each potential test item proceeds through a review process to ensure a valid, reliable, and defensible examination. Each item is reviewed and revised by a committee of Subject Matter Experts (SME) which is lead by the testing agency. At this point, the cognitive behavior classification, the item content, and the source utilized are verified. Next, each item is reviewed for its ease of reading, relevance, and the technical quality. Along with this review the keyed response is verified by the SME group.

This process continues as each test item is professionally edited by an editor or an editorial staff using style and readability guidelines. Along with this review each test item goes through a bias sensitivity review. This review is conducted by the testing agency which focuses to eliminate potentially culturally biased or offensive words, phrases, and situations.

After all of these reviews are performed and each item is approved, then the Examination Committee and testing agency are assembled together to read and select the test items for a specific test according to the percentages derived from the test specification content or subject areas. Currently, this is done before each examination administration which is twice per year. Next, the testing agency assembles the Level 1 and the Level 2 examinations for the specified test date. After the administration of the test, the testing agency conducts a preliminary item analysis, and then they gather certain members of the Examination Committee to review the preliminary item analysis and candidate comments to determine the final scoring for the examination.

Appendix A, Qualitative Evidence in the Item Development Stage

Type of Evidence	**Activity to be Performed**	**Resulting Evidence**
Item Content Definition	Role delineation, job-task analysis; practice analysis	Documentation of the methods to cite item content
Test Specifications	Table of specifications	Documentation of systematic link of test content to Specs.
Item-writer Training	Develop training materials; Train item-writers	Documentation of principles and sample items
Adherence to item-writing principles	Standard item-writing rules adopted	Evidence of compliance rules Documentation item review
Cognitive Behavior	Cognitive Classification System used to classify items	Documentation of Taxonomy Rationale & research reports
Item Content Verification	Content Experts Review and Judge items	Document sources and Relevance Reviews
Item Editing	Review Items and Professionally Edit	Credentials & Experience of Editors, Style Guidelines
Bias-Sensitivity Review	Bias-Sensitivity Review Policies & Procedures	Documentation of Bias-Sensitivity Review
Item Tryout & Pretesting	Pretest, Pilot Test; Item Performance	Documentation of examinee pilot test data, characteristics
Key Validation, Verification	Correctness of Keyed Response Verified by SME	Policy & Procedures for Key Verification
Test Security Plan	Security Policy Developed	Copy of Policy, Procedures; How Items are Protected

Quantitative Evidence in the Item Analysis Stage

Type of Evidence	**Activity to be Performed**	**Resulting Evidence**
Key Validation, Verification	Correctness of Keyed Response Verified by the Exam Committee	Procedures for Key Verification
Item Factor Analysis	Test Content Structure	Item Classification and Test Specifications Connected
Trace Line and Distractor Functioning	Plot Item Response Patterns	Item Difficulty, Discrimination, Distractor Frequencies.
Differential Item Functioning (DIF)	Deviations Among Selected Groups	Biases in Testing

Revision History:

Last Revision: 11/15/2013

Version #: 2