MEPevals

Validating the MEP Evals engineering evaluations using a content validation approach

June, 2026

Adam Schwartz
MEP Evals, LLC
202-643-6249
www.mepevals.com
admin@mepevals.com

Table of Contents

1.0 Executive Summary

MEP Evals has created evaluation tools for the purpose of improving the hiring process of engineers within the building construction industry by providing technical evaluation. These tools are intended to be used by various types of organizations, including design firms, contractors, and recruiters, to determine the level of job-related technical knowledge an engineer has. The purpose of this study is to demonstrate industry-accepted content validation of the evaluations MEP Evals has developed.

2.0 Introduction

The building construction industry is a highly technical field in which mechanical and electrical design engineers practice. Historically, the hiring process for engineers by design firms and contractors has consisted of two parts: review of the resume and standard oral interviews, whether by phone, in person, or both.

Technical evaluation of job-related industry knowledge has not been a common part of the interview and hiring process. When evaluation has occurred, it has usually been a random small set of oral questions asked by hiring managers.

Up until this point, there has not existed a validated set of technical questions which focus on essential industry knowledge for engineers in these disciplines. The technical evaluations developed by MEP Evals consist of three evaluations for each discipline (mechanical and electrical) representing different ranges of industry experience: 1-3 Years, 3-5 Years, and 5-8 years. The goal of each evaluation is to assess the technical knowledge of an individual taking the evaluation appropriate to their experience level.

This study will discuss why content validation is important, how the evaluation content has been validated, and will discuss the results of the validation process. The purpose behind performing this study is to correlate performance on these evaluations to job performance by demonstrating that the content of the evaluations has been reviewed by industry practitioners.

3.0 Validity

3.1 Assessment Validity

Before discussing the methodology utilized by MEP Evals to validate the evaluations, we must first define validity. The *Standards for Educational and Psychological Testing* (American Educational Research Association et al., 1999) (the *Standards)* define validity as:

> "The degree to which accumulated evidence and theory support specific interpretation of test scores entailed by proposed uses of a test." "Validation" is defined as: "The process through which the validity of the proposed interpretation of test scores is investigated."

The validity of an assessment is the degree to which it measures what it is supposed to measure. Validity is a relative concept, not an all-or-nothing idea.

Since we endeavor to have users utilize the scores of our evaluations as a tool in the

prediction of job-related knowledge of their engineers, we must ensure that the content of the evaluations is valid.

3.2 Content Validity

Content validity is a non-statistical type of validity that involves "the systematic examination of the test content to determine whether it covers a representative sample of the behavior domain to be measured" (Anastasi & Urbina, 1997 p.114).

The Equal Employment Opportunity Commission *Uniform Guidelines on Employee Selection Procedures* (the Guidelines) recognize content validity as an acceptable path to determine validity, and state:

“Evidence of the validity of a test or other selection procedure by a content validity study should consist of data showing that the content of the selection procedure is representative of important aspects of performance on the job for which the candidates are to be evaluated.” (Section 5B)

The *Guidelines* also state:

“In the case of a selection procedure measuring a knowledge, the knowledge being measured should be operationally defined as that body of learned information which is used in and is a necessary prerequisite for observable aspects of work behavior of the job...For any selection procedure measuring a knowledge, skill, or ability the user should show that (a) the selection procedure measures and is a representative sample of that knowledge, skill, or ability; and (b) that knowledge, skill, or ability is used in and is a necessary prerequisite to performance of critical or important work behavior(s).” (Section 14C)

MEP Evals' evaluations are intended to test an engineer's job-related technical knowledge given their experience range (behavior domain). Therefore the questions must a) be relevant to the engineering discipline being evaluated, b) must be appropriate for the age range of the exam in which they are placed, and c) must cover a representative sample of required job knowledge.

3.3 Content Validity as a Sufficient Measure

The Principles for the Validation and Use of Personnel Selection Procedures (Society of Industrial and Organizational Psychology) suggest that validity evidence can be provided to users based solely on test content. (Principles, 2003, p. 5) When “planning the validation effort...the design of the study can take many forms such as single local studies (Principles, 2003, p. 8).

In addition, “Content validity is demonstrated to the extent that the content of the assessment process reflects the important performance domains of the job. It is a validation process that can address the constraints faced by many organizations. It is a practical approach to validation. It prescribes an approach to developing the assessment process based on a study of the job. Validity is thus built into the assessment procedures. Assessment methods based on use of this strategy are usually statistically related to job performance.” [Sproule, 2009]

4.0 Content Validation Methodology

Question #79 of the *Guidelines* questions and answers states: “What is required to show the content validity of a test of a job knowledge?” The EEOC response states:

“A. There must be a defined, well recognized body of information, and knowledge of the information must be prerequisite to performance of the required work behaviors. The work behavior(s) to which each knowledge is related should be identified on an item-by-item basis. The test should fairly sample the information that is actually used by the employee on the job, so that the level of difficulty of the test items should correspond to the level of difficulty of the knowledge as used in the work behavior.”

MEP Evals has pursued two different paths to help demonstrate content validity.

4.1 A Generic Feedback Approach

The first approach was to have incumbents - people currently employed as engineers in the disciplines tested at the appropriate experience range, take an evaluation and provide any feedback that has come to mind. This free-form approach was designed to elicit feedback on question applicability and ambiguity. The results of this approach were better-worded questions and general agreement that the questions were job-related. This approach alone is not sufficient to suggest content validity. However, it provided more focused questions to use as the basis of the content validation process.

4.2 Lawshe's Approach

To establish content validity, MEP Evals elected to use a popular methodology developed by C. H. Lawshe, published in the work: *A quantitative approach to content validity*, 1975. This approach involves gauging agreement among subject matter experts (SMEs) as to how essential a particular item (question) is. Is the skill or knowledge measured by this item 'essential,' 'useful, but not essential,' or 'not necessary' to the performance of the job? In our case: is the knowledge tested on this exam essential to being able to understand and design the engineering systems that are part of the job?

Lawshe argued that if more than half of the SMEs indicate that an item is essential, that item then has at least some content validity.

4.2.1Subject Matter Experts

Lawshe (1975) later found that the best results are obtained by forming a content evaluation panel that is half comprised of incumbents and half comprised of supervisors. Incumbents have been selected to match the engineering discipline and experience range of the test they are judging.

Supervisors have been selected to meet the following criteria:

1. Licensed professional engineer in the discipline they are judging
2. Minimum of 15 years experience in the field
3. Experience in interviewing and hiring engineers

Foxcroft, Paterson, le Roux & Herbst (2004, p.49) note that by using a panel of experts to review the test specifications and the selection of items the content validity of a test can be improved. The experts will be able to review the items and comment on whether the items cover a representative sample of the behavior domain.

4.2.2 Validation Procedure

MEP Evals validated the evaluation questions using a combination of subject matter experts and incumbents. For a question to be included in an evaluation, at least half of the reviewers must have judged the knowledge tested by the question “essential.” Each SME and incumbent was asked to evaluate the following statement for each question:

> I would classify the knowledge tested for by this question to be a) essential, b) useful, but not essential, c) not necessary to understanding and designing the engineering systems which are part of the job.

Each subject matter expert was also asked to comment, in written answer form, what essential knowledge was missing from the evaluation. This survey information was used to develop additional questions to improve the breath of the evaluation and to ensure a broad and fair sampling of information topics.

After the initial results were obtained, items that did not meet the 50% threshold were replaced with new items that met the threshold, or were removed entirely. Items that were replaced were re-examined by the same panel of reviewers, and only included if they met the 50% threshold.

5.0 Conclusion

The process of content validation following the methodology described above has yielded six evaluations (3 mechanical, 3 electrical) with high content validity. Performance on these evaluations may, to a reasonable degree, be interpreted to demonstrate knowledge in the job performance domain. That is to say, the score obtained on the exam may be said to correlate with the knowledge of the engineer in their discipline at their level of experience.

6.0 Moving Forward

The MEP Evals evaluations will undergo continued validation as new questions are developed and the tests evolve over time. Each new question that is selected for use in an exam will be selected as “essential” by at least 50% of the reviewers who are asked to review them.