Chapter XVII

Evaluating Content-Management Systems for Online Learning Programs

Deborah L. Schnipke, Virtual Psychometrics, LLC, USA
Kirk Becker, University of Illinois and Promissor, USA
James S. Masters, University of North Carolina at Greensboro and Promissor, USA

Abstract

Creating quality assessments typically requires the involvement of many people who require access to the item and test information, which is stored in repositories called item banks or, more appropriately, content-management systems, since they store many kinds of content used in the test development process. This chapter discusses the types of options that are available in content-management systems and provides guidance about how to evaluate whether different content-management systems will meet an organization’s test development and delivery needs. This chapter focuses on online, fully Internet-enabled applications, since those applications have the most features.
Introduction

Everyone is assessed and evaluated throughout his or her entire life, both formally and informally. Some of these assessments are in the form of tests—either written (e.g., multiple-choice or essay tests) or performance-based (e.g., the driving portion of a driver’s license exam). Because everyone has been assessed many times, it sometimes seems that creating a test is routine and simple—write some items, package them into a test, administer the test, and give a score based on the performance. When tests are used to make important decisions about the test takers’ future (e.g., whether they can work in the field they have trained for), it is imperative that the test be psychometrically sound.

That is, the scores must be meaningful: the scores should be accurate (i.e., be reliable) and their interpretations should be backed by evidence and theory that supports the proposed uses of the test (i.e., be valid; AERA, APA, & NCME, 1999). Ensuring the psychometrics integrity of a test is much more complex than the simple steps mentioned above. Following the general approach of major test publishers (such as Educational Testing Service or ACT, Inc.), the steps for developing a psychometrically sound test might be:

1. Carefully determine the content domain of the test (e.g., focus groups with subject matter experts, job task analysis, etc.).
2. Develop a detailed test blueprint to map the test to the domain of interest.
3. Write items to match the test blueprint (e.g., in terms of specific content areas, item types, conventions).
4. Review the items for technical and editorial accuracy.
5. Field test the items to gather statistical evidence about the items.
6. If a pass/fail or similar decision will be made with the scores, determine those cutscores using psychometrically accepted methods.
7. Build the operational test forms using the field-tested items such that the test scores will be reliable, valid, and fair.
8. Publish the test for delivery.
9. Administer the test and score the test takers.
10. Monitor the test results.
11. Field test new items for future forms.
12. Refresh or replace operational test forms periodically to prevent items from being used for too long; these forms must be statistically comparable to previous forms.

Before desktop computers and the Internet were widely available, the test development process used by major test publishers was rather tedious by today’s standards. For example, one method of keeping track of item revisions and statistics was to keep every item on a 4x6” index card with statistics placed on labels on the back. As the item was
Related Content

A Case Study of Instructional Delivery Formats
Joan Aitken (2009). Handbook of Research on Assessment Technologies, Methods, and Applications in Higher Education (pp. 185-204).
www.irma-international.org/chapter/case-study-instructional-delivery-formats/19671/

Design of an AC Conductivity Measurement Setup for Sensor Materials Characterization
www.irma-international.org/article/design-of-an-ac-conductivity-measurement-setup-for-sensor-materials-characterization/126905/

Measuring and Analyzing Power Quality in Electric Traction Systems
www.irma-international.org/article/measuring-analyzing-power-quality-electric/80249/

Wavelet and Unsupervised Learning Techniques for Experimental Biomedical Data Processing
www.irma-international.org/article/wavelet-unsupervised-learning-techniques-experimental/78330/

Online Assessment in a Teacher Education Program
www.irma-international.org/chapter/online-assessment-teacher-education-program/27705/