

Proficiency, Testing and Best Practices

Dr. Karl Kapp, Professor of Instructional Technology, Bloomsburg University

Dr. Carrie McKeague, Chief Learning Officer, Kaplan EduNeering

October 2007

Abstract

A knowledgeable workforce is the cornerstone of an organization's compliance, health and safety effort. It is also a defining factor in its ability to meet performance, efficiency and risk management goals. Many organizations lack the ability to determine how knowledgeable their workforce is.

Testing is a critical tool for organizations to identify and improve the proficiency of their workforces, but tests are frequently misunderstood, misused or poorly executed.

This paper reviews the latest thinking and best practices for testing to ensure a knowledgeable, productive workforce.



White Paper

Background

A knowledgeable workforce is a company's best defense against the risks of regulatory noncompliance, compromised productivity, and workplace accidents and injuries. Despite the billions of dollars committed by businesses across the country to employee training, few organizations can determine the overall proficiency of their workforces or the knowledge of any individual within that workforce.

Knowledge is not something that can be determined by asking an individual a handful of questions. That determination requires a systematic testing approach that accommodates the performance needs of the organization and the knowledge needs of individuals to perform their job functions.

Tests are often misunderstood, misused or poorly executed, but carefully constructed testing can provide results for employers and employees:

- Aggregated test results give employers a picture of how capable their workforces are in meeting health, safety and compliance objectives. That picture, in turn, can serve as a “map” for developing a response that focuses additional learning resources where and when they are needed;
- The process of taking a test can help to educate the employee, since tests focus the learner's attention on the knowledge needed to be successful. Carefully constructed feedback to an incorrect answer can create a teachable moment, providing the right insight and knowledge to help the learner encode the correct answer and comprehend the presented knowledge;
- Although few employees enjoy tests, they frequently gain a sense of satisfaction from passing the test. That satisfaction often motivates them to do as well as possible and promotes an appreciation for the importance of the newly learned material as it relates to their job.

Practical Business Uses

Testing results can be used by business to achieve practical benefits in their recruitment, operations and compliance efforts. Among the most beneficial uses of test results:

- **Compliance:** Regulatory compliance invariably requires documented training, whether under federal agencies such as US Food and Drug Administration or in response to company-specific settlement requirements such as Corporate Integrity Agreements. In most cases, regulated organizations must demonstrate not only that they have provided training, but those learners have understood the information presented and agree to apply it. Documented testing is a key factor in demonstrating compliance or mitigating risk in the event of unforeseen events such as accidents.

- **Corporate Culture:** Companies have committed billions of dollars to create and nurture corporate cultures of quality, safety and integrity. Reinforcing the commitment and policies of companies to their employees is a key factor in successfully building and maintaining a positive corporate culture. Test results may indicate departments or locations with inadequate understanding of key subject matter, including Standards of Operation, ethics or safety. By identifying those areas, managers can mitigate potential issues and promote a more consistent culture throughout the organization.
- **Knowledge Gaps:** Gaps in corporate culture often are accompanied by serious knowledge gaps in such critical areas as safety. Testing results can quickly identify these problems, enabling managers to respond quickly to knowledge gaps that pose significant risk to the organization.
- **Subcontractor Training:** Ensuring the competency of subcontracted personnel is a growing issue, as the sponsor is ultimately responsible for the performance, compliance and safety actions of those personnel. Testing can be used to confirm that the employees of subcontractors have been properly trained prior to beginning work, and are qualified to provide the required services.
- **Remedial and Refresher Training:** An effective use of testing – whether automatically triggered by the learning management system or initiated by managers – is targeted remedial training for individuals who have demonstrated inadequate understanding of subject matter. In addition, refresher training has emerged as a key compliance component. Baseline tests showing each individual’s knowledge level enable refresher training to be conducted effectively and cost-efficiently.
- **Targeted Training:** Learners frequently complain that training is too difficult or too easy, too shallow or too extensive. In response to these issues, many learners simply tune out, either because of boredom or because they feel unable to succeed. Testing results provide managers with a clear chart of each employee’s knowledge level, enabling the targeting of training and helping to minimize the potential for bored or overwhelmed employees .
- **Proprietary Knowledge:** Testing can help measure understanding of proprietary knowledge. It can be particularly useful, for example, to determine the knowledge of sales personnel about a new product or piece of equipment. Similarly, in heavily regulated environments, such as FDA-regulated industries, demonstrated knowledge about the use and potential adverse effects of proprietary products is fundamental to regulatory compliance and risk management.

In heavily regulated environments, such as FDA-regulated industries, demonstrated knowledge about the use and potential adverse effects of proprietary products is fundamental to regulatory compliance and risk management.

Testing for Mastery

“Mastery Testing” is defined as the process of ensuring that the learner can demonstrate specified proficiency on any individual topic before proceeding to the next topic. With this validated approach, the learner must “master” the topic material rather than simply be exposed to it.

The distinction between an employee mastering a topic and simply receiving training has taken on greater significance in recent years. Regulatory compliance under agencies including the US Occupational Safety and Health Administration (OSHA), the US Food and Drug Administration requires that employees possess the necessary knowledge to perform their jobs compliantly – not merely that training has been provided by the employer. Mastery Testing demonstrates comprehension of identified topics.

Unfortunately, most people understand testing to be norm-referenced tests (NRT) like the SAT or MCAT. Unlike Mastery Testing, these norm-referenced tests compare individuals taking the tests to one another⁽²⁾. A normal distribution or bell-shaped curve is the typical result of an NRT. While people tend to cluster in the middle range, there will still be a relatively wide spread of scores. The distribution of scores produced by NRTs allows the comparison of one test-taker with another – a comparison that is at the root of tests such as SAT and MCAT.

NRTs are not an effective means for ensuring that every individual within an organization has the required competency or that everyone knows how to comply with relevant rules and regulations. To achieve that assurance, a criterion-referenced test or Mastery Test is required. In a Mastery Test environment, the performance of each test-taker is assessed without regard to the performance of any other individual taking the test. The test-taker is measured against specific criteria. Whereas interpretation of the NRT depends on the shape of the curve, interpretation of the Mastery Test requires no curve at all, since scores are analyzed according to individual competencies rather than their relationship to other test results⁽³⁾.

There is no limit to the number of people who can succeed in a Mastery Test. In fact, in most Mastery Test environments, the majority of learners are clustered at the high end of the scale because instruction given to test-takers in anticipation of the test typically parallel the competencies tested. Learners are given information they need to know to meet the criteria included in the test.

Mastery Tests are given when an organization wants to assess each individual’s specific skills, knowledge or competency. Not all Mastery Tests are equal, however. Unless properly designed, Mastery Tests are vulnerable to the same issues of invalid results as are other tests.

In a Mastery Test environment, the performance of each test-taker is assessed without regard to the performance of any other individual taking the test.

Ensuring Valid Tests

In psychometric terms, “validity” means that a test measures what it is supposed to measure. Ensuring a valid test requires that questions and the correct responses link directly to the course objectives. For Mastery Testing, two types of validity apply: face validity and content validity.

Face validity, according to Shrock and Coscarelli in *Criterion-Referenced Test Development*⁽⁴⁾, is best understood from the perspective of the person taking the test. A test has face validity if it appears to the test-takers to measure what it purports to measure. The legitimate purpose of face validity is to win acceptance among the individuals taking the test. This is a critical condition, especially for tests with significant and highly visible consequences. To ensure that tests have face validity, the test items should be taken from the content itself and written to maximize their fidelity with the objectives being tested.

Testing the New Generation

The changing workforce has direct implications for the tests used to determine the knowledge level of employees.

This new generation of employees, frequently called “Gamers,” Gen Y or the net-generation has grown up playing video games, surfing the internet and manipulating electronic gadgets and absorbing the life and learning lessons from that experience. Unlike the workforce they are replacing, Gamers are accustomed to instant feedback, learning from their mistakes, and applying the concept of “failing forward fast.” Failing forward fast means that learners gain incremental knowledge from repeated failure. A mistake in a video game is not the end result, but another opportunity to figure out a problem and try something new⁽¹⁾.

This same approach and philosophy is beginning to see application in Mastery Testing. Learners are given tests earlier in the learning process so they can self-identify what they know and what they need to know. This awareness, in turn, helps them to focus on areas of weakness. When an answer is marked as incorrect, the learner works to determine the correct answer and how to determine the correct answer when confronted with the same or similar questions in the future. This act of cognitive process encodes the knowledge and information into the mind of the learner. The act of taking the test, getting an answer wrong and being confronted by the question again is a process that solidifies the knowledge in the mind of the learner.

An important aspect of the “Gamer approach to learning” is the expectation that a test is not the terminal objective of learning. Rather, it is seen as only a part of the overall learning process that does not end with a test but includes a test as part of the complete educational process. This is an important distinction of the new workforce. Research indicates that repetition and spaced practice are two areas that can help reinforce the encoding and recall of knowledge – two critical factors when attempting to bring a workforce to a desired level of mastery.

The different approaches and perspectives of the gamer generation impose new requirements on training if it is to be effective. Instructional design has evolved to meet these new requirements, which include interactivity and gaming elements. New training courses are dynamic, incorporating techniques such as simulation, heightened graphics use, and music. Most important, however, is the format of the training programs. To capture the imagination and attention of Gamers, courses should offer multiple opportunities to succeed – something that the previous generation of employees would not have accepted or understood. Training for Gamers is most effective when it is short in duration, focusing on one issue rather than many.

Content validity, which is required for any Mastery Test to be valid, is distinct from face validity. The main distinction between face validity and content validity is that face validity is determined by non-experts and content validity is determined by experts. A test possesses content validity when recognized experts have verified that the test measures what it is supposed to measure.

Best practices for content validity include the active participation of both subject-matter expert and instructional designer. The subject-matter expert examines the content, helps establish the objectives, and scrutinizes the test items to ensure accurate content. The instructional designer ensures that questions match the objectives of the course and that the questions are written in the appropriate format.

A common misconception is that a test has to cover every topic in the course to be valid.

A common misconception is that a test has to cover every topic in the course to be valid. In fact, a test need not cover all content in any given course to have content validity; instead, it must cover a representative sample of the content. The content tested for mastery must be contained in the course and must be relevant to the job of the learner – and that relevance must be determined by a qualified subject-matter expert.

Content validity is not determined by the opinions of test-takers, but by the consensus of subject-matter experts. If a group of experts had reviewed the questions and content of the test and had deemed the question to be important to the job, the test can be considered to have content validity. The test overall would be valid since content validity is the more rigorous standard of the two types of validity.

Another misconception centers on the need for a consistent response to a wide range of questions for the test to be valid. For example, if a group of learners took the same test and every learner failed every test item, is the test invalid? The answer is “no” if the test has been vetted by a group of experts and has been shown to have content validity. The fact that every test-taker answers a question incorrectly – or correctly – is not an indication of the test’s validity.

A test is most useful if it can accurately distinguish among people who know the correct answers, people who think they know the correct answers, and people who are guessing.

A test is most useful if it can accurately distinguish among people who know the correct answers, people who think they know the correct answers, and people who are guessing. The ability of a test to identify these distinctions is known as the “Difficulty Index.” The Difficulty Index indicates the number of people who answered a test item correctly. Typically, this statistic is expressed as a decimal. For example, .70 means that 70% of the people tak-

Task	Objective	Performance Criteria
Inspect a vial and accept or reject based on three critical criteria.	When given a vial, the employee will be able to correctly accept or reject the vial based on three evaluation criteria within four minutes with 100% accuracy.	Given an image of a vial that can be rotated on the computer screen, the learner will be required to indicate if the vial should be accepted or rejected with 100% accuracy within four minutes which are counted down on the screen.

ing the test have answered the item correctly. The range is from .00, signifying that nobody answered the item correctly, to 1.00, which indicates that everyone answered the item correctly.

In Mastery Tests, it is not uncommon for a large number of people to answer the question correctly since the courses are designed specifically to promote mastery of the content. The Difficulty Index on many items would be in the 0.80 – 1.00 range. If learners are not gaining mastery, there may be a problem with the instruction since mastery is the goal. Although the concept of the Difficulty Index should be kept in mind when designing the test, the ease or difficulty of test items for any individual or group of learners is not an indication of the validity of the item or the test.

Creating Effective Testing Items

The best-known and most versatile testing format is the multiple-choice format. Multiple-choice questions can assess learner knowledge at four levels: memorization, comprehension, application and analysis.

A well-constructed and effective multiple-choice question includes a question (known as the stem) and a list of possible answers to complete the stem. The incorrect answers from which a learner can choose are called distracters or foils, while the correct answer is usually called the key. Typically, multiple-choice questions contain three distracters and one correct response, for a total of four possible answers.

While any test item format is important for creating an effective test item, the most critical element is linking the test item to a specific learning objective. This process ensures content validity and a fair and unbiased test item. It also links tasks the learner needs to perform to an assessment of his or her ability to perform the task. The most effective method of ensuring a valid test is to ensure that the questions and the desired responses link back to the course objectives. The three-column method for creating items is frequently identified as a Best Practice.

Using this type of chart, you can distinguish between different types of learning when you create your assessment item. You can tell if you are testing at one level but the objective is at another. For example, you may want to distinguish between a learner’s ability to memorize a concept and the learner’s ability to apply the concept.

An inappropriate question for the above task would be, “Identify the three criteria that lead to the rejection of a vial.” In this case, this question is asking for identification and not application of a skill. Just because a learner may know the three criteria, that doesn’t mean he or she can apply that criteria. The test question must ask the learner to apply the concept, not merely repeat the memorized criteria. Do not ask the learner to identify the three success criteria when you really want them to apply the criteria. The level at which you are testing must match the level at which you expect the learner to perform. When this happens, the questions are valid.

While any test item format is important for creating an effective test item, the most critical element is linking the test item to a specific learning objective.

Using this method, a task that needs to be learned is linked to a specific learning objective and a specific assessment item. This linked approach ensures a congruent connection among the task to be performed, the learning objective taught, and how the learner is assessed. The philosophy and discipline of this three-column approach ensures the creation of effective mastery assessment items .

Best Practices for Choosing Training/Testing Technology

Even though there are many technology options available for training and testing from many different vendors, companies should be careful about choosing technologies that offer optimal functionality. An effective training/testing solution should enable the following:

- Individual distribution of training materials based on job, employee, location, department or changing organizational structure;
- Electronic grading, eliminating the need for in-person grading while ensuring accuracy of testing;
- 24/7 monitoring, allowing managers real-time oversight of the training status of individuals, courses or curriculum;
- Rapid versioning;
- Multiple uses, allowing the system to be used for training as well as critical information control, routine corporate communications and solicitation of information from employees;
- Accessibility for multiple departments, allowing training to be coordinated throughout the organization and promoting the use of test results for multiple purposes ;
- Content validity that clearly relates the test questions to the knowledge needed;
- Accurate measurement of a workforce's knowledge through a test creation process that ensures validity and the presentation of test questions that are randomized from user to user to reduce the opportunities to share answers.

Conclusion

Testing has shown itself to offer significant benefits to organizations in managing their workforces, operations and compliance. It determines an individual's understanding of critical information, how to apply that information (knowledge) on the job, and gaps in knowledge that could impact individual or company performance. Mastery Testing is one of the most effective learning philosophies because it focuses on the individual rather than the group. It is a philosophy that encourages learners to gain the knowledge they need to ensure the health, safety and compliance of themselves and others. Most important, it is a fair and effective method for ensuring that employees receive the most benefit from the training they receive.

References

[1] Kapp, K. M. (2007) *Gadgets, Games and Gizmos for Learning: Tools for Transferring Know-How from the Boomers to the Gamers*. Pfeiffer, New York.

[2] Shrock, S. & Coscarrelli, W. (2000) *Criterion-Referenced Test Development: Technical and Legal Guidelines for Corporate Training and Certification*. Second Edition. International Society for Performance Improvement, Silver Spring, MD. Pg. 25.

[3] Shrock, S. & Coscarrelli, W. (2000) *Criterion-Referenced Test Development: Technical and Legal Guidelines for Corporate Training and Certification*. Second Edition. International Society for Performance Improvement, Silver Spring, MD.

[4] Shrock, S. & Coscarrelli, W. (2000) *Criterion-Referenced Test Development: Technical and Legal Guidelines for Corporate Training and Certification*. Second Edition. International Society for Performance Improvement, Silver Spring, MD.

[5] Borg, R. W. & Gall, M. D. (1989) *Educational Research: An Introduction*. Fifth Edition. Longman, White Plains, New York.

[6] Shrock, S. & Coscarrelli, W. (2000) *Criterion-Referenced Test Development: Technical and Legal Guidelines for Corporate Training and Certification*. Second Edition. International Society for Performance Improvement, Silver Spring, MD.

[7] Shrock, S. & Coscarrelli, W. (2000) *Criterion-Referenced Test Development: Technical and Legal Guidelines for Corporate Training and Certification*. Second Edition. International Society for Performance Improvement, Silver Spring, MD. Pg. 74-75.

[8] Shrock, S. & Coscarrelli, W. (2000) *Criterion-Referenced Test Development: Technical and Legal Guidelines for Corporate Training and Certification*. Second Edition. International Society for Performance Improvement, Silver Spring, MD. Pg. 75.

About Kaplan EduNeering

Kaplan EduNeering (www.kaplaneduneering.com), a Deloitte Technology Fast 500 Company, is part of Kaplan, Inc., a worldwide education services leader and a subsidiary of The Washington Post Company (NYSE: WPO).

EduNeering develops technology-enabled knowledge solutions for improving business performance and assuring regulatory compliance. Headquartered in Princeton, NJ, Kaplan EduNeering has offices in Houston, TX, Bloomsburg, PA and the United Kingdom.

Kaplan EduNeering's Headquarters
202 Carnegie Center
Suite 301
Princeton, NJ 08540
Tel: 609-627-5300
Fax: 609-627-5330

Kaplan EduNeering's Houston, TX (US) Office:
EduNeering
3700 Bay Area Blvd.
Suite 550
Houston, TX 77058
Tel: 281-709-0900
Fax: 281-709-0907

Kaplan EduNeering's Bloomsburg, PA (US) Office
30 East 6th Street
Bloomsburg, PA 17815
Tel: 570-387-9121
Fax: 570-387-9122

Kaplan EduNeering's Europe Office
95 Wilton Road, Suite 58
London SW1V 1BZ
England, UK
Tel: +44-203-084-7893



© 2007 Kaplan, Inc. All Rights Reserved.
WP100107