Standardized Test

By Anthony M. Wanjohi

1.0 Introduction

The earliest evidence of standardized testing based on merit comes from China during the Han dynasty. The concept of a state ruled by men of ability and virtue was an outgrowth of Confucian philosophy. The imperial examinations covered the Six Arts which included music, archery and horsemanship, arithmetic, writing, and knowledge of the rituals and ceremonies of both public and private parts. Later, the five studies (military strategies, civil law, revenue and taxation, agriculture and geography) were added to the testing.

Standardized testing was not traditionally a part of Western pedagogy; based on the sceptical and open-ended tradition of debate inherited from Ancient Greece, Western academia favored the essay. However, given the large number of school students during and after the [Industrial Revolution], open-ended assessment of all students was not viable. Moreover, the lack of a standardized process introduces a substantial source of measurement error. .

2.0 Description of Terms

2.1 Standardized test

A standardized test is one that is administered under standardized or controlled conditions that specify where, when, how, and for how long children may respond to the questions or "prompts." Standardized tests should meet acceptable standards for technical qualities in construction, administration, and use. (Goodwin and Driscoll, 1980)

Standardized tests have the qualities that provides a systematic procedure for describing behaviors, whether in terms of numbers or categories, including specified procedures for administration and scoring, test items that are derived from experience, either by experiment or observation, rather than theory, present the same tasks and require the same response modes

from all test takers, provide tables of norms to which the scores of test takers can be compared in order to ascertain their relative standing (Goodwin and Driscroll, 1980).

2.2 Validity of Test Results

Validity is the accumulation of evidence to support a specific interpretation of the test results. Since validity is a characteristic of test results, a test may have validities of varying degree, for different purposes. The concept of instructional validity relates to how well the test is aligned to state standards and classroom instructional objectives (AAC, 2003).

Langhorst (1989) states that it is evaluated by the extent to which the instrument measures what it purports to measure. Validity, therefore, is always in regard to a particular test use.

2.3 Reliability

Reliability refers to the consistency of test scores. Various methods are used to calculate and estimate reliability depending on the purpose for which the test is used. Adequate reliability depends upon a low magnitude of errors of measurement. Such errors can be introduced, for example, by inconsistencies in the performance of those being tested or variability in their interest, motivation, or other emotional and physical states, which may be unrelated to the purpose of the test. Such influences make stability in test scores of young children very problematic (Langhorst, 1989).

3.0 Procedures in Test Administration

Hills (1992) stated that individually administered standardized tests can serve important purposes when they are selected wisely and used as intended by the test developers, and their results are not over generalized or misinterpreted. School personnel must not, however, limit assessment to such tests or allow the tests to dominate assessment of young children. He quotes it later that "The primary strengths of standardized tests, if they are properly designed and properly used, are that they can eliminate biases in assessment of individual children and that they provide data that can be aggregated, permitting comparisons of groups to a standard.

The goal in administration of standardized tests is to obtain the best possible sample of student behavior under standard conditions. Strict adherence to the test manual's guidelines for

administration and scoring is an absolute necessity. The manual usually provides instructions and test items to be read to the student.

Test administration includes carefully following standard procedures so that the test is used in the manner specified by the test developers. The goal of Test administration is to obtain the best possible sample of students' behavior under standard conditions. Strict adherence to the test manuals that provide instructions to the test and guidelines that are highlighted is absolutely necessary. The test administrator should ensure that test takers work within conditions that maximize opportunity for optimum performance (JCTP, 2002).

3.1 Before administration of standardized test

According to AAC (2003), before administration of standardized test, it is important that relevant persons (students, teachers, administrators) are informed about the standard testing procedures, purposes of the test, the kinds of tasks involved, the method of administration, the scoring and reporting and have sufficient practice experiences prior to the test to include practice as needed. Further they should provide information on how to operate equipment for computer-administered tests and practice in response to tasks, sufficiently trained in their responsibilities and the administration procedures for the test, have a chance to review test materials and administration sites and procedures prior to the time for testing to ensure standardized conditions and appropriate responses to any irregularities that may occur, arrange for appropriate modifications of testing materials and procedures in order to accommodate test takers with special needs and have a clear understanding of their rights and responsibilities.

3.2 What should be done during administration of test

American Counseling Association (1997) observes that the following should be done during administration of test:

The testing environment: Testing environment includes seating, work surfaces, lighting, room temperature, freedom from distractions and psychological climate that are conducive to the best possible performance of the examinees. The tester should make every effort to create a comfortable testing environment with adequate space for sitting arrangements, comfortable seats

and desks for sitting and writing. There should also be adequate lighting and ventilation to ensure comfortableness in the test room.

Test administrators: Test administrators should be sufficiently trained to be able to establish and maintain uniform conditions and observe the conduct of students when large groups of individuals are tested; They should also follow the instructions in the test manual; demonstrate verbal clarity; use verbatim directions; adhere to verbatim directions; follow exact sequence and timing; and use materials that are identical to those specified by the test publisher.

Procedure/Test protocol: A systematic and objective procedure should be put in place for observing and recording environmental, health, emotional factors, or other elements that may invalidate test performance and results; deviations from prescribed test administration procedures, including information on test accommodations for individuals with special needs, are recorded.

Security: The security of test materials and computer-administered testing software is protected, ensuring that only individuals with a legitimate need for access to the materials/software are able to obtain such access and that steps to eliminate the possibility of breaches in test security and copyright protection are respected.

3.3 What should be done after administration of test

Testing ends when all the scheduled activities are completed and the time allotted is exhausted, or the student is no longer able to work efficiently due to fatigue or loss of concentration. At the end of the session, the tester should thank the students for their cooperation and explain what will happen next. There might be further testing another day, meeting of the students and tester to go over results. The administrator should collect and inventory all secure test materials and immediately report any breaches in test security and include notes on any problems, irregularities, and accommodations in the test records (ACA, 1999).

These precepts represent the basic process for all standardized tests and assessments. Some situations may add steps or modify some of these to provide the best testing milieu possible.

Test Scoring: Scoring is a critical step that transforms the students test response into comparative data. Accurate measurement necessitates adequate procedures for scoring the

responses of test takers. Scoring procedures should be audited as necessary to ensure consistency and accuracy of application. These procedures include careful implementations or monitoring the standard of scoring procedures, human judgment, use of rubrics that clearly specify the criteria for scoring, monitoring constant consistency and providing a method for checking the accuracy of scores when accuracy is challenged by test takers (M.E. 1999).

Communicating test results: Before communication of test results takes place, a solid foundation and preparation is necessary. That foundation includes knowledge of test interpretation and an understanding of the particular test being used, as provided by the test manual. Conveying test results with language that the test taker, parents, teachers, clients, or general public can understand is one of the key elements in helping others understand the meaning of the test results. When reporting group results, the information needs to be supplemented with background information that can help explain the results with cautions about misinterpretations. The test user should indicate how the test results can be and should not be interpreted (NCREL, 2003).

4.0 Conclusion

A standardized test also known as a norm-referenced test is a type of a formal test that has a structured procedure with specific guide lines for administration, scoring and interpretation of the results. The results of the tests are used to as a measure to determine eligibility of special education and identification of strengths and weaknesses in school learning. The results also help professionals to plan instructions by identifying curriculum areas in which students fail to perform as well as their peers. Proper test use resides with the test user (ACA, 1997) the counselor and educator. Qualified test users understand the measurement characteristics necessary to select good standardized tests, administer the tests according to specified procedures, assure accurate scoring, accurately interpret test scores for individuals and groups, and ensure productive applications of the results. This document provides guidelines for using tests responsibly with students and clients.

References

Association for Assessment in Counseling (AAC) (2003). *Responsibilities of Users of Standardized Tests.*

Prepared by the Association for Assessment in Counseling (AAC).

American Counseling Association (APA)(1997). Code *of ethics and standards of practice.* Alexandria, VA:

Goodwin, W. L., & Driscoll, L. A. (1980). *Handbook for measurement and evaluation in early childhood education*. San Francisco: Jossey-Bass.

Hills, T. W. (1992). *Reaching potentials through appropriate assessment*. In S. Bredekamp & T. Rosegrant (Eds.), *Reaching potentials: Appropriate curriculum and assessment for young children* (Vol. 1, pp. 43-63). Washington, DC: National Association for the Education of Young Children.

Joint Committee on Testing Practices (2002). *Code of Fair Testing Practices in Education.*

Langhorst, B. H. (1989). *Assessment in early childhood education: A consumer's guide*. Portland, OR: Northwest Regional Educational Laboratory.

Measurement in Education (1999).*Standards for educational and psychological testing.*Washington, DC: American Educational Research Association.

Suggested Citation in APA

Wanjohi,A.M.(2010). *Standardized Test. KENPRO Publications.*Available online at http://www.kenpro.org/papers/standardized-test.htm

Access PDF

Click to access printable format for PDF Download