

# HISTORICAL REVIEW ON ACHIEVEMENT TESTING: INTERNATIONAL PERSPECTIVE

*Dr. Iweka, Fidel. E.G.*

## **Abstract**

This paper is a historical review on achievement testing, and international perspective. It is an x-ray of the historical development of achievement testing from the primitive societies to the present advanced societies, from its informal form to the present formal form. The paper also highlighted the stages in the trend of formal assessment as greatly influenced world-wide by social, economic and ideological goal of the society up to the present stage of continuous assessment and greater teacher involvement in certification. It examines some of the central rationales driving test-based accountability models also referred to as standards-based reform and discusses the necessity of shifting to data-integrated decision-making that affectively balances different forms of student assessment for accountability purposes.

## **Introduction**

Assessment has been an integral part of man right from the beginning. In primitive societies, they were informal but in present advanced societies they are both formal and informal. It has been argued that formal assessment practices which emerged from formal education is one of the clearest indices of the relationship between school and society since they provide for communication between the two (Broad foot, 2000).

## **International Perspective**

In Europe, formal educational assessment started with what is called public examination in only about a century ago. But before this, achievement test had already been used in pre-nineteenth century society of imperial china as the civil service entrance examination (Broad foot, 2000). In Europe, it was economic and social forces that led to the initial institution, of various types of 'quality control' mechanisms in education and to the massive development of educational achievement test over the last hundred years. Bell and Broant, (2001), noted that achievement tests on emergence were quickly institutionalized in the national educational system and ever long became invested with a legitimacy as unassailable, as the headmasters gown.

The new demand the capitalist economy was making on educational provision soon reflected in the development of assessment procedures as a tool by which schooling was made more responsive to the needs of the economy. To reflect the changes that were taking place, in the social role of education and issues that have come to dominate education in modern industrial societies ever since, led to the development of new assessment practices.

Dubois (1999), identified four interrelated themes in the development of assessment practices which reflects clearly the change that were taking place in the social role of education and issues that have come to dominate education in modern industrial societies as "competence", "content" competition", and "control".

These internationally, brought about one of the most marked trends in assessment, which concerns the postponement of selection. Example of such is three division of secondary education, which has been replaced by a common school for all children up to the end of compulsory schooling.

Post second world war school registration was a system of "sponsored" mobility with common education terminating well before the statutory school leaving age and children being identified apparently on merit, as future elite members, technocrats at any early age.

The "grundeschute" was introduced in Europe for all pupils in Sweden in 1946, a model already in USA and soviet union which was there perhaps because of a long standing pre-occupation in these countries with education as a means of creating equality of opportunity and national unity in contrast to the more single minded European association of secondary schooling, with academic excellence.

Social and ideological forces caused a more comprehensive reform referred to as "postponement of differentiation". Thus, Dore (2001), argued that the system gave raise to 'qualification inflection'-more pupils achieved qualification at each level of schooling and hence these

qualifications having less and less value as selection instrument.

The fourth stage in the trend of assessment is that reported by Neave (1999), which posits a law of educational development in which assessment for all constitutes what might be termed "the universalization of extended secondary education". He reported that examination at the end of compulsory schooling has typically given place to a regular series of standardized tests and have to allow teachers to gauge their pupil's progress in relation to that of the country as a whole.

This points to the new emphasis worldwide for increasing teacher involvement in certification.

This new trend towards teacher-based certification is as a result of an increasing level of public trust and confidence in schools and teachers. For a long time, teachers have become very reliable persons in ranking pupils accurately on both specific abilities and non-cognitive qualities, (Mangoran 2000). Examinations are of crucial importance for selection and they must be invested with as much apparent objectivity as possible so that the results and the failure which they imply for many candidates are accepted.

The trends of educational assessment as could be seen has moved from informal to formal assessment. Assessment has developed and changed according to societies' role expected of it and it has been greatly influenced worldwide by social economic and ideological goal of the society up to the present stage of continuous assessment and greater teacher involvement in certification.

### **Review of Educational Assessment in West Africa**

In pre-ancient, pre-Christian times, they used oral test to assess abilities. The first oral test was administered in china in 115 B.C, it was designed to assess abilities in civil service examinations.

By the middle of 8<sup>th</sup> century, its major weakness has been found. It was neither reliable nor objective. Although it is still used today in assessing abilities. Another form of assessment has taken precedence and this is the written test now widely used.

West African countries just like other countries in the world over used oral test to assess abilities. This was mostly informal. However, by 1800 AD, written examination was introduced with the advent of western education since then, two broad categories of examinations have developed.

The two examinations are the internal test administered by the teachers in their schools and the external examinations, (Onuche and Akeju, 2001).

Nwana (2002:184), reported that by the recommendations of a Royal Commission (the Jeffrey Commission) set up by the British Colonial Administration established in 1952, the council was charged with the responsibility of conducting school examinations in West Africa countries of Ghana, Sierra Leone, Gambia and Nigeria.

It was empowered to conduct such examinations and to award certificate that will be of equal standard with certificates of examining authorities in United Kingdom. In 1966, the Test Development and Research Unit (TEDU), an organ of West African Examinations Council was established. TEDU develops standardized tests for the council. In 1992 the International Center for Evaluation was established at the university of Ibadan to among other functions, train evaluators, provide talents and personnel for the evaluation of specific projects.

Since then, the Nigerian government has established other examination bodies charged with different functions. The Joint Admission and Matriculation Board (JAMB) was established in 1978 to cater for admission into Nigerian universities.

The Nigerian government in 1982 adopted a new method of assessment, the continuous assessment, as opposed to the former terminal examination at the end of each educational level, National Policy on Education, (1981).

The trend of educational assessment followed world trend and is equally influenced by social, economic and ideological factors. Beeby2000: 67 stated:

To the extent third world countries seek to evaluate western industrial society and seek to engage in the world economy, they are likely to come to model their education system in response to economic social and political influences very similar to those affecting advanced industrial societies and go through similar stages of development.

### **Current Trend In Achievement Testing- International Perspective**

Governments and senior educational policy-makers in much of the Western World have

explicitly endorsed the philosophy that test-based accountability represents the best way to spur improvements in schools. By holding teachers accountable for external test performance, educators are compelled, to improve their classroom practice, thereby raising the educational performance of all students and also narrowing the gap between low and high achieving students.

### **Assumptions Guiding Standards-Based Reform:**

There are a variety of interrelated rationales that drive standards based reform movements around the world. Rather than discuss the merits and limitations of each of these assumptions, the researcher focuses on three key prepositions that underline the increased use of standardized testing to drive school improvement.

1. High-stakes testing is the most effective policy-lever to improve teaching and learning.
2. A focus on getting back-to-basics by testing key curriculum areas such as language arts, mathematics and potential science, is the most effective way to gauge educational quality and to improve schools; and
3. Standardized test results are the most reliable and valid form of data for planning and improvement purpose. Closer scrutiny of each of these rationales suggests the need to re-examine the basis of contemporary standards-based reform initiatives.

### **Policy-Lever Rationale:**

The chief rationale driving standards-based reform is that high-stakes testing is the most effective lever to improve teaching and learning within schools. Regional and National Testing targets provide benchmarks to gauge educational progress and provide necessary stimulus to spur school improvement.

This assertion, however, is not supported by the empirical literature, in fact, four large reviews on the impact of formative assessment have demonstrated that these practices double the spread of students learning (Black & Williams, 1998; Crooks, 1988; Kluger and De Nisi, 1996; Natriello 1987).

Improved formative assessment helps struggling students more than others and thus reduces the achievement gap while still raising achievement overall.

Thus, formative assessment, not summative standardized testing, provides teachers with the most salient data to inform their practice and improve student learning and achievement.

Although it is difficult to draw comparisons across nations with diverse student populations, current trends suggest that the intensification of external testing for accountability purposes has not led to significant improvements in achievement.

For example, countries with the most intensified external testing programmes, such as England and United States, have not fared well on international measures such as the trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS), and Programme for International Student Assessment (PISA).

Ironically, countries such as Finland, which possess no centrally administered, testing measures, seem to be leading the pack in the previous international assessments. These findings dispel the notion that countries which emphasize external testing are more likely to raise student's achievement.

Back-to-basics-rationale, another important rationale used to support standards-based reform is that a focus on key curriculum areas such as language arts and mathematics is essential for school improvement. Other non-core subject areas such as music, physical education, visual arts, etc. are too costly and difficult to assess and may actually provide a distraction from what should be a focus on getting back-to-basics. The end result is that schools inevitably become oriented toward dual-focused (language arts, mathematics, science) versus multi-focused in their attention to the mandated curriculum. When one considers the long-standing criticism that most standardized measures focus on low-level thinking skills, the prospect of developing a critical mass of students ready to spur innovations in the knowledge economy is greatly diminished.

The advent and wide spread use of standardized testing measures is partly a response to the perception that teachers' classroom assessment data are more subjective and less reliable and valid than standardized measures. In truth, part of this assertion is currently unassailable, classroom-based evaluation judgment can vary significantly between teachers. Nevertheless, this seems unlikely to change in systems that put resources and excessive energy in external programmes.

The underlying message for teachers working within such systems is that their judgments are inconsequential for significant decision making purposes and they do not necessarily have to improve their current assessment literacy and understanding of the principles and practices of sound assessment (Stiggins, 2004). If classroom assessment is to have more robust reliability, then practicing teachers require ongoing professional development opportunities to develop their assessment expertise in a context that values those judgments (Volant, 2006).

### **Moving Forward with Data- Integrated Decision-Making**

Date-driven decision-making has become an increasingly important concept with the multitude of assess available to administrators and teachers in contemporary schools. Administrators and teachers<sup>1</sup> utilization of assessment data for planning purposes can take various forms. The first and lowest level response, typically involves the examination of large scales results in isolation at other forms of student data (Volante, 2007). Here teachers and administrators make adjustments to teaching and planning on the basis of general test scores in particular subject areas. This approach often results in large - scale assessment driving, rather than informing school improvement.

The second level is similar to the first, with the exception that large — scale assessment data is disaggregated for particular student groups (special needs students, English as a second-language students, distinct ability groups, etc). Although the second level shows greater sophistication, it still suffers from the same methodological problem, large-scale assessment drives teaching / planning decisions without the benefit of other forms of student information.

The third level, and highest level involves the integration of disaggregated large - scale assessment results with other forms of student assessment information. Educators at the third level make teaching/planning decisions based on multiple and at times, contradictory forms of student assessment information. These educators are able to examine student's assessment information across arrange of subject areas, some of which cannot be assessed in a standardized fashion.

### **Conclusion**

In his review of the relationship between standardized testing and school improvement. Raudenbush (2005) concluded that measures of mean proficiency and other statistical analyses of test data alone do not reveal direct evidence of the quality of school practice. Policy — makers can facilitate this correction by enhancing the salience of classroom - based assessment data for accountability purposes. Interestingly, research has shown how classroom assessment data has been successfully integrated for standards - based accountability purposes. Thus, classroom assessment should be an integral part of an overall accountability framework so isolated, from other forms of student assessment information.

Jurisdictions must develop data - integrated decision - making systems that utilize multiple forms of student assessment data to inform, classroom, school, and district planning. In doing so, they provide the public with a more comprehensive approach to assessing school improvement and are well positioned to respond to the complex demands of the 21<sup>st</sup> century.

### **Reference**

- Black, P. and William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phidelta kappan*, 80 (2), 139-148.
- Black, P. and William, D. (2005). Lessons from around the world: How policies, politics and cultures constrain and afford assessment practices, politics and cultures constrain and afford assessment practices. *The curriculum Journal* 16 (2), 249 - 261.
- Crooks, T.J. (1988). The impact of classroom evaluation practices on students. *Review of Educational research*, 58 (4), 438-481.
- Eckstain, Max A. (1992). *Examinations: comparative and international student*. New York: Pergamon press.