# The Role of Assessment in Teaching and Learning.

**By**

**Peter Odor; Francis Ojo; Matthew Audu**
**Joint Admissions and Matriculation Board (JAMB), Nigeria.**

Paper presented at the 21st Annual Conference of AEAA held at Cape Town, South Africa, 25th to 29th August, 2003.

# The Role of Assessment in Teaching and Learning.

**By**

**Peter Odor; Francis Ojo; Matthew Audu**
**Joint Admissions and Matriculation Board (JAMB), Nigeria.**

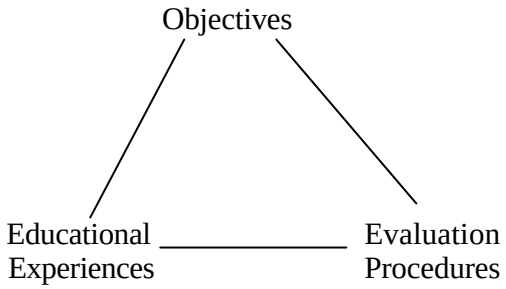
## Abstract

The paper notes the traditional view on the relationship between testing and teaching/learning. It also explores emerging thoughts, which recognize the potential of testing in influencing teacher and student behaviour in terms of how teaching and learning take place. While recognizing the dangers inherent in the misuse of testing in this regard, the paper considers the emerging thoughts as welcome and worth implementing in a developing society by external examination agencies. The paper further expatiates on the rationale and methods of implementing the new approach and points out the implications to examination bodies in particular and the educational system in general, of implementing this new role of testing.

## 1.0 Traditional View of the Place of Assessment in Teaching and Learning.

Traditionally, testing is seen as a means of determining the effectiveness of teaching and learning. As such, it becomes an end-of-instruction endeavour and a means of providing feedback to the teaching and learning process. Mehrens & Lehmann (1991) observed that historically, testing is regarded as an activity that is thought about only after instruction is over. The normal order of things is that the kind of changes desired of the pupils (i.e. the objectives) are determined; the educational experiences or instruction necessary for realizing the objectives are identified; the instructional materials and strategies and the staff development programs (where needed) are developed. They further cited Furst (1958)

schematic representation of the relationship among instruction, objectives and evaluation as follows:




Thus, by the traditional view of testing, tentative, preliminary objectives determine the instructional procedures and the methods used to evaluate both educational experiences and objectives. At the same time, evaluation and educational experiences help clarify the objectives, and the learning experiences help determine the evaluative procedure to be used. Moreover, the results of evaluation provide feedback on the effectiveness of the teaching experience and ultimately, on the attainability of the objectives for each student (Mehrens & Lehmann, 1991).

In sum, by the traditional view, assessment is dependent on or is driven by the objectives and the educational experiences.

## 2.0 Emerging View of the Place of Assessment in Teaching and Learning.

The emerging view of testing varies from the traditional view. By the traditional view, while testing may remain an end-of-instruction endeavour, by the new approach, a test may not always be keyed wholesale to the specific objectives of the instruction or depend wholly on the content of instruction. Depending on the purpose to which the resulting scores may be put, the content/focus/emphasis of a test may drive instruction/learning rather than be driven by instruction.

For example, where the information about student achievement is strictly for internal school use, using tests that are directly linked to specified educational objectives and instruction is quite plausible and is in consonance with the central tenet of the evaluation movement (Airasian & Madaus, 1983). Also, in the summative evaluation context, if the intention is just to determine whether a curriculum/instruction achieves the stated objectives, the traditional view becomes predominant. However, it is not always the case that the objectives and educational experiences form the focus of assessment. In summative evaluation, a situation may arise where the intention is to determine both the explicit and implicit (pay-off) objectives consequent on the teaching and learning. In this case, the assessment has to go beyond the explicit objectives. The findings therefrom would definitely influence future teaching and learning. Mehrens (1984) described as undebatable the fact that, in evaluating the adequacy of the curriculum one needs to go beyond determining whether those objectives taught have been achieved. He notes that, when we want to make inferences about instructional quality, a close match between instruction and the test would be required. But for most other inferences we may wish to make, we may not necessarily want a perfect match between specific instructional objectives and the test. Certainly, for curricular decision making, we wish to know about something more general. Cronbach (1963) made this very clear when he stated:

> **"In course evaluation, we need not be much concerned about making measuring instruments fit the curriculum. However startling this declaration may seem, and however contrary to the principles of evaluation for other purposes, this must be our position if we want to know what changes a course produces in the pupil. An ideal evaluation might include measures of all the types of proficiency that might reasonably be desired in the area in question, not just the selected outcomes to which this curriculum directs substantial attention". (p.680).**

**Cronbach (1971) further stated:**

> **"The recommendation that the evaluation battery be comprehensive seems to run counter to the concept that an educational test should measure what has been taught. And students think a test "unfair" when it asks about topics not covered in the course. One can agree that it is unjust to let the fate of an individual be determined by a test for which, through no fault of his own, he is ill prepared. But this only illustrates once more how a test valid for one decision can be invalid for another. "Though it is unfair to judge the quality of a teacher's work by a test that does not fit the course of study he was directed to follow, that test may be a fair basis for judging the curriculum. If teacher plus course-of-study have left the pupil ignorant of contemporary literature, this is a significant fact about the adequacy of his education" (p.460).**

While agreeing on the need to broaden the basis of assessment beyond the explicit (objectives and educational experiences), educators such as Green (1983) and Mehrens (1984) buttressed this point further when they suggest that the emphasis of a test should extend to more useful skills. In the words of Green (1983):

> **"if the students have learned fundamental skills and knowledge and understand it, they will be able to answer many questions dealing with the materials not directly taught … generalized skills and understandings do develop… Since all the specifics can never be taught …This development is highly desirable and tests. should try to assess it. This can only be done by having items that ask about content not directly taught" (p.6).**

**Mehrens (1984) stated:**

> **"Frequently both educators and members of the general public wish to make inferences to a broad domain, whether or not that whole domain has been taught. …If we wish to infer whether students can interpret graphs in real life, we wish to make a more general inference than one tied directly to some teacher's instruction. Teachers should also want to make such inferences. They should care about transfer of knowledge even if they only wish to evaluate the effectiveness of their own instruction. If parents wish to infer how well their children will do in another school next year they need to infer from a general domain, not from the perhaps narrow and idiosyncratic domain of a single teacher's objectives. The whole basis**

**behind giving various standardized achievement test batteries is that tests covering fairly general domains provide valuable information." (p.11).**

What comes through from the foregoing, is that the purpose that a test generally will serve may force a departure from the traditional approach. Experts in educational assessment such as Bloom (1969); Airasian (1979); Popham (1987), etc, believed that the purpose to which the resulting test score will be put also has a significant effect. Where the decision being made at least in part, from test score is one with important consequences, the test is described as a "high-stakes" test. The stronger and more influential the intended use of test results, the stronger and more influential the tests become in the affected person's eyes Bloom (1969); Airasian (1979) also observed that by making the tests important, one encourages a variety of responses and alterations in practice to conform to the nature and content of the tests.

Specifically, if the sanctions or rewards associated with test performance are important, then the test can have enormous the influence over what is taught, how it is taught, what is learned and how it is learned (Madaus & Mcdonagh, 1979; Madaus & MacNamara, 1970; Madaus & Airasian, 1977; Madaus & Greaney, 1982).

**The term measurement-driven instruction MDI** is used when high stakes test influence teaching and learning. According to Popham (1987), measurement-driven instruction occurs when a high-stakes test of educational achievement, because of the important contingencies associated with the students' performance, influences the instructional program that prepares students for the test. He described how such high stakes tests operate when he stated:

**"Few educators would dispute the claim that these sorts of high-stakes tests markedly influence the nature of instructional programs. Whether they are concerned about their own self-esteem or their students well-being, teachers clearly want students to perform well on such tests. Accordingly, teachers tend to focus a**

> **significant portion of their instructional activities on the knowledge and skills assessed by such tests. A high-stakes test of educational achievement, then, serves as a powerful curricular magnet. Those who deny the instructional influence of high-stakes tests have not spent much time in public school classrooms recently" (p.680)**

The concept of measurement-driven instruction has its critics. Some of the criticisms are that such tests can: corrupt and destroy the very process they are intended to monitor or improve (Campbell, 1975); shift control over the curriculum and teaching to the agency that controls the test and this fact can change the structure and emphasis of an educational system (Airasian, Madaus & Pedulla, 1979); lead to a devaluing, undervaluing, or ignoring of those aspects of schooling that are not included on them, thereby narrowing and restricting the curriculum (Holmes, 1911); distort teaching until it becomes a "game of mechanical contrivance" aimed at getting pupils to pass (Sutherland, 1971); and lead to coaching, which in turn can lead to a distortion of the traits which the tests were originally to measure (Linn, 1983; Madaus & MacNamara, 1970). In the words of Airasian and Madaus (1983), MDI tests have the power to force a match between testing and instruction, but at a cost that may make the match meaningless.

However, a number of educators see high stakes tests as necessary evil, one that serves a host of policy-oriented purposes which include: assessing educational quality; providing evidence of school and programme effectiveness. Airasian & Madaus (1983) commented that like it or not, we live in a normative world, a world in which comparisons among individuals and institutions are not only wanted, but also actively sought. Taxpayers, parents, administrators and legislators – for a variety purposes – wish to make comparisons between pupils, teachers and schools. For the most part, these publics are unable or unwilling to accept testimonials about the goodness of education from those whom they perceive to have vested interest in that goodness, namely teachers and administrators. These publics want some external, objective assurances of the

quality of teaching and learning, and so they turn, virtually by default, to the one index they believe provides the external objective evidence they seek; standardized achievement tests.

In their view, it is not necessarily bad that tests influence what goes on in classrooms if these tests reflect what has been agreed upon as achievable goals for classrooms. If there is a middle ground in the debate over the benefits and liabilities of extra-classroom accountability, it likely will be found in the understanding that goals, instruction and testing are integrally related and therefore must all be considered in framing accountability schemes.

Popham (1987) emphasized that MDI can be a potent force for educational improvement. To him the tests can serve as a powerful curricula magnet directing instructional efforts to important objectives. In his words:

> **"If properly conceived and implemented, measurement-driven instruction currently constitutes the most cost-effective way of improving the quality of public education… It is a strategy for instructional improvement that … educators should consider seriously" (p.679).**

The words of Airasian (1988), depict the level of confidence that can be reposed on the new approach in his observation that despite existing criticisms most advocates and critics seem to agree that MDI can work and that high-stakes tests can and will drive instruction.

## 3.0 Conditions that Maximize the Effects of High-Stakes Tests.

According to Popham (1987), to ensure the effectiveness of measurement-driven instruction, a high stakes test must be characterized by the following:

(i) must assess genuinely defensible skills and knowledge (i.e. non-trivial knowledge and skills including higher level behaviours, must be measured);

(ii) the skills and knowledge tested must be clearly described;

(iii) focus of the test must be on a reasonable number of important skills or knowledge targets which must be sufficiently general to subsume enabling skills and content;

(iv) the content should be subjected to intense scrutiny by all concerned clienteles, so that the tests measure truly worthwhile content and

(v) the skills or knowledge to be tested must be conceptualized in such a way, that teachers can use the target skills and knowledge to design instructional sequences – thus the test can function as catalyst to improve instruction.

In addition, the following conditions for achieving maximal efficiency are worthy of note:

(i) the greatest impact on instruction will occur when high-stakes and high standards tests are present (Popham, 1987; Airasian, 1988);

(ii) in general, fewer pupils will have their instruction driven by a high-stakes, low standard testing programme than by a high-stakes, higher standard programme;

(iii) the more course-specific the test content, the more easily instructional responses can be localized in the curriculum sequence and the more likely that the test will drive instruction and conversely and

(iv) the more the test content is spread across grades and the curriculum, the more difficult it will be to organize instructional responses and the less the impact the test will have on instruction.

## 4.0 The Realities in the Context of External Examinations in Nigeria.

The reality is that, the examinations conducted by most of the external examination bodies are inherently "high-stakes" tests as they have important impact on the lives of the students, the parents, the teachers and administrators. As such, there is no gainsaying that the external examinations do have the potential to influence educational practice. The examination bodies therefore, need to embrace and capitalize on the emerging reality and thoughts concerning the examinations they conduct and make necessary adjustments where they have not done so. Major considerations may relate to the perceptions of the examination bodies of: the role of the school in educational/societal transformation; the role of the examination bodies in providing direction for effective societal transformation; and the focus of testing within the context of MDI.

## 4.1 Role of the School in Societal Transformation

To carry out the improvement related functions to society, examination bodies must take into consideration the assertions of Airasian (1987), about the place of schools in societal transformation. In his view:

(i) There is a growing recognition that among all social institutions, the schools are the most feasible agency in which to locate many social initiation and reform efforts.

(ii) Schools are at once a conservative social institution, in so far as their self-impetus for change is not great, and a battlefield on which the agenda of varied social and political groups compete for attention and implementation.

(iii) Just as we might perceive ourselves to be ill or our bodies to be malfunctioning and seek diagnosis and cure, so too, do we now diagnose the malfunctioning in our society and seek ways to change it.

In otherwords, while schools are important change/reform agents, they are often not the initiators of the change/reform. Other agencies initiate these changes which are attempts at solving identified social needs. The needs assessment is carried out by all social service agents including examination agencies, with the ultimate aim of achieving a better society through the instrumentality of the school.

## 4.2 Role of External Examination Bodies in Providing Direction and Effecting Change.

With respect to the role of providing direction and effecting change, the examination bodies should liken their roles to those of the judiciary (vis-à-vis the Executive and the Legislature). The judiciary interprets the laws of the legislature. This interpretation often takes a number of factors into consideration. The factors include among other things, the constitution of the country, the interest of the nation as an entity, the mood of the nation, the psychological impact of their decision, and the reaction of the international community to their decision. In otherwords, interpretation of the law is usually not done in a vacuum. Serious effort is made to achieve some balance on a number of obvious mediating factors while trying to achieve the obvious – equity, justice and fairness. This places the nation in a good stead among the committee of nations. Stability within and

portraying the nation in good light appear to be moderating factors. Once the decisions are made the other arms of government adjust appropriately.

In the same vein, the educational assessment bodies through the testing program(s) are the interpreters of the curriculum and the educational objectives. The curriculum developer can be likened to the Legislature. The products of the interpretations and analyses of the curriculum and the objectives form the basic input into the testing program(s). In doing this, interest should go beyond the explicit to the implicit. In addition, the educational, scientific, social, economic and other forms of developmental needs of the country should be taken into consideration.

The emphasis of the activity and what the agency ultimately settles on depends on which of two main schools of thoughts regarding the function of the school/education. Tyler (1949: 35/36) articulated this in the following philosophical questions. **"should the educated man adjust to society, should he accept the social order as it is, or should he attempt to improve the society in which he lives?"** However these questions may be answered, the answer will in turn affect the emphasis of measurement. If the belief is that the primary function of the school (and by implication testing) is to teach people to adjust to society, it will strongly emphasize: obedience to the present authorities; loyalty to the present norms and traditions; and skills in carrying on the present techniques of life. Whereas, if the emphasis is on the revolutionary function of the school, concern will be more on critical analysis, ability to meet new problems, independence and self-direction, freedom, and self-discipline. The nature of the philosophy the testing agency adopts will affect the slant of the educational tests developed and used. The inherent nature of the high-stakes tests would suggest a tilt toward the revolutionary function of the school.

## 4.3 Focus of Testing in the Context of High-Stakes Tests

Popham (1987) advised that, because high stakes tests will surely function as instructional magnets, the challenges (to test developers) is to fashion those tests so that they become a potent force for educational improvement. To achieve this, there is the need to reconceive the relationship between measurement and instruction, so that tests are employed as vehicles of instructional clarification and skills clarification/standard definition.

### 4.3.1 Instructional Clarification

Instructional clarification relates to analysis of the curriculum content and this involves two major considerations: national interests/needs and developments in the international scene. In each, interest may not be on re-writing the curriculum content but on identifying essential gaps that need to be filled. With respect to national needs, the initial test development effort may involve an in-depth interpretation and analysis of existing curricular content with a view to identifying non-trivial but related areas of curriculum content and highlighting same in the examinations bearing in mind existing national needs. The creative genius of the assessment can thus be measured in terms of its ability to identify important gaps and sensitize the educational system enough to effect necessary improvement-oriented reforms.

In-depth interpretation and analysis of existing curriculum content may also be carried out with the intent of identifying non-trivial but related areas of the curriculum content and highlighting same in the examinations bearing in mind existing state of the knowledge in the international scene. The present level of globalization requires constant peep into the outside world with a view to either solving internal problems and/or just keeping abreast with happenings out there. In education, we must move with the times, we must learn to reject the irrelevant to make room for the relevant. We must not be so presumptuous as to think that if

it was good enough for us, then it must be good enough for our children. The world they will grow into is not the world we have lived in. We must accept that there will always be evolutionary change. We must always be ready to examine what is new and we must not be afraid to use experience gained elsewhere.

Comparative education study and adaptation is not an activity of curriculum development which occurs at specified time periods only. Test developers need to take up the challenge of the time and deal with the issues on here-and-now basis rather than allow the nation to "wait" into the future for these to feature in the classrooms. This wait may ultimately prove dangerous, or, the implementation too late for meaningful effect. For example, the case of emergence of AIDS/SARS and the need to emphasize basic skills in hygiene, examiners should not wait until the concept of AIDS or SARS is included in the curriculum for them to be sensitive to such issues in the examinations. The resolve of the test developers should be to sensitize the school system to the emerging changes and trends as they come, rather than wait for the next three to four years when actual curriculum review may incorporate the changes. Assessment thus, has become a veritable tool for keeping the nation as system "open" (and not " closed") to happenings in its environment. It is true that open and adjusting systems do not die. The same may not be said of closed system.

### 4.3.2 Skills Clarification/Standard definition

The skills clarification/standards definition component relates to the relative emphasis of the tests on the cognitive skills: lower levels, higher levels or both. Improving the standards requires moving from the lower to the higher cognitive level. Available studies on the cognitive skills emphasized by objective tests in Nigeria point to over emphasis of lower level skills. Odor, Solanke & Azeke (1986) found that, WAEC did not go beyond the application level in all its examination of 'O' Level CESAC Biology for the period 1975-1985. Also Eze (1986) found that 'O' Level Physics questions for 1982-1985 were very few in the

areas of analysis and synthesis but non-existent in the area of evaluation. A study by WAEC (2002) listing of abstracts reported that except in Mathematics, the objective test items in all other subjects from 1997-2002, featured objective items that tested more of the lower cognitive levels than the higher cognitive levels.

Such emphasis on factual knowledge and comprehension than on higher level thinking processes portends a danger. Olson (1980) described the danger thus:

> **"Your life can be spent acquiring more and more facts and knowledge, yet the end product will be an aged and immature child still unable to think and act in a creative, effective, and productive manner, or your life can be successful, exciting and productive, when you develop your creative resources and make creative thinking an integral part of your life" (p.2)".**

Nations cannot achieve greatness by allowing their young to operate at the lowest level of thinking process. Great heights are achieved when the youths are encouraged to "dream dreams" and to transform their visions into reality (Olson, 1980). The need for higher level cognitive skills arises from our low level of economic/technological development. The high-stake tests especially the objectives tests have been shown to measure basic skills as if they are the only skills that could be measured. There is need to recognize the essence of developing high level thinking processes especially in a developing consumer-oriented society. Developing tests that deal with genuinely higher order skills will improve our perspectives and our ability to deal with issues and situations. We need to develop tests that assess more demanding skills. There is no gainsaying it that higher order questions are the foundation to acquisition of skills basic to educational, scientific and technological development. Moreover, other nations are embracing this shift towards higher order questions, and Nigeria cannot afford to be left out. For example, in January this year, China embarked on a major revolution to overhaul its educational system in order to meet up with modern challenges. Learning by rote system is being gradually replaced with emphasis on producing enough **innovative thinkers.** This year alone about 1.5

billion will be spent providing on-the-job training for teachers. (This DAY Newspaper. Monday ,May26,2003, p.80).

In apparent acceptance of this "new deal", Salim (1997), in his capacity as the Registrar of JAMB declared:

…it is not JAMB's task to train candidates to the level required. It is, however, JAMB's task to ensure that the country has to get good products as much as possible in line with the goals of both the educational and technological development of the country.

There are sets of universal standards. And JAMB has to ensure that the products of our higher institutions are at par or even better than what can be universally accepted as basic educational qualification or educational conducts expected of either a university, polytechnic or college of education graduate anywhere.

We, there fore, cannot lower the standard to make it easier for more people to pass because that would make JAMB to be churning out candidates that will not be useful to anybody in the country. As you know, this will ultimately tarnish the image of the educational system of the country.

The only thing we can do is to ask pupils to prepare themselves adequately. But as one of the custodians of the educational hopes and aspirations of this country, we cannot afford to lower our standards to bring about greater quantity without the desired quality. Quality is what counts. (Thisday Newspaper, Thursday January 2, 1997 p.26)

## 5.0 Implications of the Emerging View of Testing

A prominent feature of the emerging view of testing is the emphasis on measuring higher level cognitive skills and abilities – skills of problem solving, application, thinking and reasoning. Implicit in this desire to measure higher order cognitive skills and abilities are a number of instructional complexities which have to be recognized and taken into consideration in order to achieve the desired outcomes of instruction – the high level cognitive skills.

The instructional complexities relate to two major issues:

(a) availability of sufficient instructional knowledge and techniques to enable teachers to teach most pupils the desired objectives.

(b) what the tests that measure higher level skills will actually measure.

Airasian (1988) put the issues in proper perspective when he observed that though implicit in MDI is the view of instruction as a mere technical, problem, readily solvable by application of existing standard procedures, this not necessarily the case. He observed that the key instructional tasks involved in teaching higher level behaviours which are to; represent new problems in terms that the students recognize; select from their repertoire of information rules, principles and experiences that will help them solve the problem; and apply the knowledge to solving the problems, are not simple processes to teach. He admits that though we are able to train pupils to solve well-structured problems in narrowly defined knowledge domains, we know little about teaching pupils to solve more general higher level problems. He however, provides some clues on the way out when he noted that various reviews indicate that teaching higher level behaviours is different in many ways from teaching lower level rote behaviours. For example, higher level behaviours:

- take longer to teach;

- develop gradually over times, not by the quick association and closure that often occurs in teaching rote materials;

- call for more indirect, less structured instructional environments and methods; and

- involves providing opportunities for pupil self-discovery.

With respect to what to measure, Airasian (1988) observed that no clear body of knowledge exists regarding the conceptualization of high level behaviours such as reasoning, critical thinking, and application, nor are there well-validated instructional strategies to teach such higher level processes.

Thus, efforts at using high-stakes test to influence instruction need to address the two problems – those of identifying valid instructional techniques and teaching same to teachers and developing credible definitions for the higher order cognitive behaviours.

## 5.1 Implication to the Examination Bodies

Examination Bodies need to realize that, testing, especially the high-stakes tests which they conduct is more of specialist job, in which experts in testing and also in curriculum should be involved. The present realities where participation of measurement experts in test development and administration is the exception rather than the rule should be avoided. A way out may be to evolve rigorous in-house training exercise in Measurement and Evaluation for the curriculum experts already in the employment of the agencies. In addition to this ad-hoc approach to

solving the problem, conscious effort should be made to send staff with relevant background for higher degrees in the area of Measurement and Evaluation.

Another issue that comes up, is that of developing credible definition for higher order cognitive behaviours. Achieving a common language is basic to advancement in any given field. Our inability, to date, to operate beyond the lower level cognitive skill may stem from the fact that, we may not know much about the actual meaning of each of the higher level cognitive processes as pointed out by Airasian (1988). We must admit that the whole area of educational assessment and the technicalities, especially the language of it is new and a little bit foreign. The reality is that, achieving credible definitions for the respective cognitive processes is not beyond our capability. We as a group may have to sit down and drawing from our different backgrounds/disciplines and experiences, begin to ask ourselves, what for example, 'analysis' means in operational terms; what behaviours are overt indication of analysis as a human endeavour.

Towards this definition, a workshop may be organized by the committee of examination bodies in Nigeria. This workshop should draw participants from various disciplines like sociology, psychology, philosophy, education and the respective subject areas. The aim of drawing participants from the various disciplines, is that each will bring to the fore his specialist knowledge in effort to define each of the cognitive levels. It is hoped that for each cognitive level, the final product will be a definition in very simplified forms and languages. Such specific observable behaviour will then form the basis of constructing the test items and thus facilitate test development.

Another area that the testing agencies collectively and individually may have to focus on, is the development of **item writing manual**. Input into such document may form part of the training of test developers. However, each testing agency has to develop the principle that will guide a test constructor and assessor of the test items. It is in this regard that, a testing manual gains primacy.

## 5.2 Implication to the School

Airasian (1988) painted a fairly gloomy picture of implementing high-stakes tests when he observed that the key instructional tasks involved in teaching higher level behaviours are not simple to teach. He however, gave some ray of hope when he pointed out some of the research findings about teaching/learning with respect to higher level behaviours. Of particular interest are, the assertions that **higher level behaviours: take longer to teach; develop gradually over time; require indirect, less structured instructional environments and methods; and involve providing opportunities for pupil self-discovery.** These observations, collectively and individually, have implications to the teachers, school administrators and the school itself, in the effort to promote development of higher level cognitive behaviours.

For example, the knowledge that higher level behaviours take longer time to teach than the lower level cognitive behaviours and develop gradually over time should prompt the teacher and indeed, the school, to introduce the teaching of higher level behaviours from day one of the child's education. Spreading this practice over such a considerable period of time will enable the teachers/students to imbibe the pattern of teaching/learning as a way of life. In doing so, the teacher should design the in-class and out of-school assessment measures in such a way that they emphasize high level thought processes. Known educational practices have proved that awareness forces the sensitive teacher/student to seek/search for viable solutions and that practice makes perfect.

Teachers should realize that an interplay of the content of instruction, and the various ways the teacher conducts the actual classroom interactions affect the students in terms of: how they look at life; approach problems/issues; and how they deal with problems ultimately. If the teachers ask high order questions, if they use varied approaches to asking questions and demand varied approaches to answering questions, the students would definitely imbibe these behaviours. To

the extent these are essential behaviours the nation will benefit and thus, be the better for it.

It is the responsibility of the school and all stake-holders in education to ensure that the teacher's training gives him the skills, understanding and the patience required to teach these high level behaviours. With respect to instructional method and environment, the scientific or discovery method should be encouraged, where the child explores knowledge from the unknown to the known. The school on its part should provide adequate infrastructure, instructional materials and manpower to encourage the development of higher level cognitive skills.

## References


Airasian, P. W. (1979). The effects of standardized testing and test information on teachers' perceptions and practices. Paper presented at the annual meeting of the American Educational Research association, san Francisco, California, April, 1979.

Airasian, P. W. (1987). State Mandated Testing and Educational Reform: Context and Consequences. The University of Chicago Press. May, 1987.

Airasian, P. W. (1988). Measurement Driven Instruction Education, Washington, Writer, 1988.

Airasian, P. W & Madaus, & G. F. (1983). Linking Testing and Construction: Policy issues. Journal of Educational Measurement Vol; No. 2. (1983).

Airasian, P.W; Madaus, G.F; Pedulla, J. J. (1979). Minimal competency testing, Englewood Cliffs, N.J: Educational Technology Publications.

Bloom, B. S. (1969). Some theoretical issues relating to educational evaluation. In R. Tyler (ed). Educational evaluation; New roles, new means. The sixty-eighth Yearbook for the National Society for the study of Education, Part II. Chicago, Illinois: National Society for the Study of Education, pp 26-30.

Campbell, D. T (1975). Assessing the impact of planned social change. In G.M. Lyons (ed.), Social research and public policies: the Dartmouth/OECD conference. Honover, N.H. Dartmouth College, Public Affairs Centre.

Cronbach, Lee J.(1963). Course Improvement through Evaluation, Teachers College Record, Vol. 64.

Cronbach, Lee J. (1971). Test Validation. In R. Thorndike. Educational Measurement ($2^{nd}$ ed.) (pp.443-507). Washington, Dc. American Council on Education.

Eze, A.E (1986). Evaluation through multiple-choice question: An appraisal of WAEC type in SC/GCE (O/L) Physics. In O. O. Bello (ed), $27^{th}$ Annual Conference Proceedings of the Science Teachers' Association of Nigeria (pp.278-282). Ibadan: STAN.

Furst, E. J. (1958). Constructing Evaluation Instruments. New York, Mckary.

Green, D. R. (1983). Content Validity of Standardized Achievement Tests and Test Curriculum Overlap. In D. Wanous (Chair). National vs. local test and curriculum: Inferences to which domain and why. Symposium conducted at the annual meeting of the National Council on Measurement in Education. Montreal.

Holmes, E.G.A (1911). What is and what might be: A study of education in general and elementary in particular London: Constable & Co.

Linn, R. L (1983). Curricular Validity: convincing the courts that it was taught without precluding the possibility of measuring it. In G.F Madaus (ed.). The courts, validity and minimum competency testing. Hingham, M.A; Kluwer – Nijhoff Publishing, Boston.

Madaus, G.F. (Ed.) 1983). The courts, validity and minimum competency testing. Hingham, M.A: Kluwer – Nijhoff Publishing, Boston.

Madaus G. F. & Airasian, P.W (1977). Issues in evaluation student outcomes in competency-based graduation programs. Journal of Research and Development in Education. 10(3), 79-91.

Madaus, G. F & Greaney V. (1982). Competency testing: a case study of the Irish Primary Certificate Examination. Occasional Paper Series. National Consortium on Testing, Cambridge, MA.

Madaus, G. F & Macnamara, J (1970). Public Examinations: a study of the Irish primary certificate examination. Dublin: St. Patrick's College, Educational Research Centre.

Madaus, G. F & Mcdonagh, J (1979). Minimum Competency Testing: Unexamined assumptions and unexplored negative outcomes. In R. Lennon (Ed.), Impactive changes in measurement: New directions for testing and measurement, 3, 1-14.

Mehrenn, W.A (1984). National tests and local curriculum: Match or mismatch?. Educational Measurement Issues and Practice, 3(33), 9-15.

Mehrens, W.A & Lehmann, I. J. (1991). Measurement and Evaluation in Education and Psychology. (4th Edition) New York: Holt, Rinehart & Winston. P.8

Odor P. I.; Solanke, O. O; Azeke, & T. O. (1986). An appraisal of the objective test in WASC 'O' Level DESA biology: In O. O. Bello (Ed) 27th Annual Conference Proceedings of the Science Teachers' Association of Nigeria (pp 199-204): Ibadan: STAN.

Olson, R. W (1980). The Art of Creative Thinking. Harper & Row, Publishers. New York.

Popham, W. J (1987). The Merits of Measurement-Driven Instruction. Phi Delta Kappan. 1987.

Sutherland, G. (1971). Elementary education in the nineteenth century. London, London Historical association.

Tyler, R. A (1949) Basic principles of curriculum and instruction. The University of Chicago Press, Chicago.

West African Examination Council (2002). The Taxonomy of SSCE Questions by Cognitive Domain: (1997-2000)., In List of 2002 abstracts of Research Reports from the Research Division of WAEC. (RPA/2002)

This Day Newspaper (Nigeria). Thursday, January 2, 1997, p.26

ThisDay Newspaper (Nigeria), Monday, May 26, 2003 p.80.