

GLOBAL LEARNER ASSESSMENT CHALLENGES AND SOME INNOVATIVE RESEARCHES FOR OVERCOMING THEM

Professor P.M.C. Ogomaka
Imo State University,
Owerri.

Abstract

The world, as many say, is now a global village, so most education theories principles, models, policies and practices are common to most national/international education systems. Learner assessment theories, models, principles, policies and practices are not exempt from the acclaimed "commonness". Ironically the strength of global resources in education has not successfully overcome the pull and push, gloves in hands factors that impede/militate against the use/application or implementation of these common theories/models, principle and policies. The global learner assessment challenges include but not limited to: the lack of and willingness to use instruments that appraise learners in all domains of learning outcomes; lack and non-use of learner assessment instruments that have content validity; the inequities inherent in the use of assessment instruments, the scoring of learners responses to such instruments and the uses of results obtained there from. Some researchers/innovators including this presenter have; developed some instruments, thought out and illustrated, some procedures for overcoming the challenges. Recommendations and suggestions are made as regards the ways forward.

Keywords: Assessment, Challenges, Global, Innovation, Learner/Learning and Research.

The term global implies "objects", issues or experiences that concern the entire globe, the world we live in. Something, an "object" or experience is said to be global if it is; common to, found everywhere all over or spread all through the earth we live on. The global learner could mean everybody who learns (human being and any other individual animal species that does learn). Learning could be a process or a product. It could be seen as the process of acquiring knowledge, skills and dispositions (KSDs). According to one of the age old definitions, learning implies a relatively

permanent change in behavior which is strictly not ascribable to physical growth or tissue damage. The learner and learning thus far are too omnibus. The learner and learning concerned in this work, are those in formal education institutions and disciplines/subjects studied therein. Learner assessment in this context implies the assessment of how and what the learner has learned in school. Assessment is synonymous with evaluation. In its narrowest sense, learner assessment/evaluation is a process or mechanism for determining how much the learner or group learners has

learned/acquired from a (curricular) content to which the learner or group learners has been exposed. Learner assessment/evaluation implies much more than the above. In its broad sense, learner assessment or evaluation entails ascertaining:

- i. The quality or comparative extent of functioning of some completely or improveable endowed/innate psychological, psychosocial or physical abilities,
- ii. The comparative extent of acquired knowledge, skills and/or dispositions as a result of exposure to some curricular content;
- iii. The style and comparative rate of acquiring the knowledge, skills and dispositions (ksd) and
- iv. The best option (advance, change, restrategies or terminate) with regard to programme/instruction, specialization/career for the learner.

Thus far the global learner and what the global learner assessment entails have been explained, albeit briefly, what are the global learner assessment challenges? First challenges, in this context, imply situations, circumstances, factors or tasks that:

- i. defy (known methods, theories, principles, practices or explanations);
- ii. as it were, call for a fight or repositioning

- iii. require/demand special effort on extra-work/new or modified approaches to overcome and
- iv. impede, inhibit or stall the achievement of set goals/objectives of some enterprise. (Ogomaka 2012). The enterprise here is global learner assessment. These challenges in specific terms are many and will be taken on one by one later in this presentation.

As implied in the general explanation of challenges, they are situations/circumstances that impede, inhibit or defy the achievement of goals and objectives so to overcome the challenges, the situations/circumstances /factors have to be removed or changed. The changing must not be haphazard. They have to be studied and purposefully removed so innovative research is needed to overcome the challenges one by one.

Innovation implies a purposeful and planned change. In this context the plan must be for the better. Research is systematic investigation/study geared toward: extending/advancing the frontiers of KSDs and/or solving a felt problem (even at times only "exposing" the felt problem). Indeed innovative research is very much needed for overcoming the challenges.

Global Learner Assessment Goals and objectives

Perhaps before handling the global learner assessment challenges, it is pertinent to outline the goals and objectives of learner assessment /evaluation which the challenges impede,

inhibit or stall their achievement. These goals and objectives do include but not limited to:

i. ascertaining the comparative standing of each learner/group of learners among universal, regional or local set of learners in intelligence, giftedness, creativity, sensory acuity, physical strength/dexterity, psychosocial, social responsiveness;

ii. ascertaining the comparative quality and amount of ksd (call them cognitive, affective, or psychomotor learning outcomes) the learner acquires/posses/exhibits after exposure to some curricular content over a period of time or at the end of a programme

iii. ascertaining areas of strength, weakness or comparative advantage of a learner or group of learners as regards some curricular frame/programme and

iv. decide on advancing, continuing, modifying or changing a programme for a learner.

To achieve the goals and objectives, scores/grades obtained mainly from evaluation instruments of course many of the learner assessment challenges are consequent upon the extent of expertise/appropriate technical know-how of the: instruments developers, users of the instruments and the users of the scores/grades obtained therefrom. Most of the times the instruments used are tests of various types/forms. The development of the instrument should be based on appropriate theory or model with respect to what the instrument is to be

used for. Also the instrument should be of quality having desirable characteristics of assessment instruments.

Theories/models upon which instrument development may be based

There are two main sets of theories/models upon which assessment instruments' development are based. One is the Classical Test Theory (CTT) and the other is the Latent Trait Theory (LTT) or the Item Response Theory (IRT). CTT is a much more earlier theory than IRT. It still influences instrument development and evaluation practices till date.

The major proposition of CTT is that any measurement of an "object" as an observed score, X say, has two components a true score, T say, and an error score, E say.

Thus $X = T + E$ where E is a random variable which could be positive or negative. In words, every observed score of an "object" is either an overestimate or an underestimate of the true score of the "object". An extension of this theory is that if in an area of study, individual A performs better than individual B and B performs better than individual C then A performs better than C.

The first part of the theory underlies the practice of Continuous Assessment and requirement of multiple reading of measures/readings in laboratories. The second part of the proposal is the basis of Test Item analyses yielding difficulty/facility and discrimination indices of test items. However the results of item analyses based on CTT have always been criticized as

being sample specific [Nworgu, (2016); Ogomaka (2016) & Ainsworth (2015)].

Furthermore the use of correlation method in computing reliability and criterion related validity indices are based on CTT.

(Ogomaka; 2016) The indices obtained from the above are also said to be sample specific (Ogomaka 2016 & Ainsworth 2015)

IRT is a more recent theory than CTT, dating back only to the 1960. IRT focuses on two major things: (i) an individual has specific latent traits (innate abilities) which the individual uses to tackle specific challenges as presented by test items, etc, each individual's specific latent trait (innate ability) could be calibrated and each could attain a level denoted, θ say. Also the level of challenge posed/presented by each item could be calibrated and if the level it attains is denoted, b say. Then in situations where: (i) θ is greater than b the individual scores the item correct; (ii) θ is less than b , the individual fails to get the item correct and (iii) θ is equal to b , the individual has fifty fifty percent chance of getting the item correct or wrong. Besides the, b -parameter or difficulty level of an item, there are the a and c parameters, and probably more, that may affect an individual responds to an items. So IRT being a probabilistic theory prescribes the probability of a testee with an ability level θ will respond correctly to: an item with difficult parameter b as $P(\theta/bi = \text{correct}) = \frac{\exp(\theta-b)}{1+\exp(\theta-b)}$

An item with parameters, b (difficulty) and a (discrimination) as $P(\theta/bi, a = \text{correct}) = \frac{\exp[a(\theta-b)]}{1+\exp[a, (\theta-b)]}$

an item with item parameters b , (difficulty) a , (discrimination) and c (pseudo guessing factor)

as $P(\theta/bi, a, c = \text{correct}) = C + (1 - C) \frac{\exp[a(\theta-b)]}{1+\exp[a, (\theta-b)]}$

The probability functions/models shown here above are the one-parameter, two-parameter and three-parameter models respectively. In the process of instrument development, these models (by the application of computer packages) are used to determine and calibrate item parameters. Items parameters so established and calibrated are believed to be non-sample dependent/sample independent (Ogomaka 2016 & Nworgu 2016).

Learner Assessment Challenges

As implied in the introduction, there are a number of learner assessment challenges. These challenges are presented under subheading as: coverage of the domains of learning outcomes; coverage of curriculum content; and inequities in testing processes and practices, and inequities in grading and classification

Coverage of Domains of Learning Outcomes

At the present level of human development or probably discussions, there are three domains of learning outcomes. These are: (i) the cognitive domain; (ii) the affective domain and (iii) the psychomotor domain as conventionally classified. The

cognitive domain has been originally categorized into (approximately) a taxonomy viz; knowledge comprehension (understanding), application analysis, synthesis and evaluation. (Tracy, 2015 & Ogomaka, 2016)

Recently, however the taxonomy is revised and presented as remembering, understanding, applying, analyzing, evaluating and creating (Tracy, 2015).

Though not part of this work but in a lighter or observer mood this new taxonomy needs to be reconsidered following these questions or comments: is creating above being evaluated? - at the end of creating, GOD found it good...; is creating as act of creativity, conclusively/exclusively learned, or an outcome of leaning... whether revised or not the challenge here is that our test items as constructed by teachers and examination bodies, do not include at all or enough items that appraise the higher order cognitive skills (applying, analyzing, synthesizing and evaluating or creating, whatever it means here). Another challenge related to the non-inclusion or very low inclusion of items that appraise the higher cognitive skills is that there are no conventions/agreements on the proportion/percentage of items that appraise each category of the cognitive skills at the various levels of our educational system.

A greater/more disturbing challenge is the very little inclusion of avenues for assessing/appraising learners' affective learning outcomes. Though a misconception/misnomer, character and learning are the stated basis for graduating

learners in higher institutions. There are records, scores, grades and perhaps classifications of learner assessment in the cognitive domain but little or none in the affective domain. In Nigeria in particular attempts at appraising the learners' affective learning outcomes at the Basic Education level are highly compartmentalized in terms of categories. Such attempts are based on armchair observation using no instruments.

The challenge involved in appraising of learners' affective learning outcomes is diversified by the lack of instruments and complicated by the thinking that the affective domain subsumes character and moral behaviors. The assessment of a learner's cognitive domain is wholistically done per subject/discipline, could the affective learning outcomes be so done?

The avenues for appraising the learners' psychomotor learning outcomes abound, there are no much disagreements in the categorization of the psychomotor domain into a taxonomy/hierarchical subdomains. The learners' learning outcomes could be appraised in most subjects/disciplines to various extents, but the much the learners are appraised are most of the times in adevertently integrated into the scores and grades awarded to the cognitive learning outcomes of the learners (writing of letters and figures, quality sound production, free hand sketches/diagrams of objects, constructions in Mathematics, Technical/Engineering drawings, wood work, metal work, making of ridges,

fences etc) could these not be scored separately in various subjects and then integrated somehow? Another challenge is that there are some psychomotor learning activities/tasks that some learners cannot participate in. how could those students be accommodated if the assessment of learners in the psychomotor domain is carried out comprehensively as done in the cognitive domain?

Coverage of Curriculum Content and Programme Objectives

The number of test items that could be drawn from a subject/broad-based curriculum content is certainly very large and impossible to be covered in a number of formal tests. Some learner assessment tests can only be meaningful if they are presented as situations, projects or practicals etc. These kind of tests are difficult to construct and are rarely used. Since it is impossible to construct all the possible tests and impossible to administer all the possible tests to a given learner/group of learners, making use of a representative sample of all the possible tests/test items is advocated. However the way and manner of getting at the representative sample of all the possible tests/test items is a big challenge (Ogomaka 2016 & Ogomaka & Osuala 2015).

Another big challenge lies in the use of tests/test items' objectives to reflect and cover the curriculum/programme's objectives.

In a question form, to what extent do tests or test items' objectives agree with and cover the concerned subject's objectives? Supposing the objectives of the

unit in mathematics on commercial Arithmetic include: (i) enabling the student/learner to buy and sell goods and services and (ii) set up a small commercial business/store; could the manner of tests/test items on addition and subtraction, multiplication and division of money, profit and loss as done/responded to in classes, assignments or examinations reflect comprehensive or cover the objectives? Does learner assessment in schools help learners to carry out tasks expected of them in their real life and places of work after school? How can assessment situations and conditions in school resemble undertaking real life task?

Challenges inherent in Testing processes and practices - As Inequities

Testing conditions, processes and practices referred to here include but not limited to: (i) timing of tests; (ii) giving test item options (iii) sitting arrangements, (iv) items arrangements (v) identifying legitimate testees etc. Inequality entails: assigning/giving/awarding some honour/task/prize/position to an individual who did not, merit/have capability for the/win/qualify for the thing respectively; assigning insufficient or too much time to individuals to undertake a task, an assignment for a due reward/price or giving equal things to unequal individuals.

Most of the times, the length of time assigned to learners for completing a test is arbitrary. Casually this researcher asked an incidental sample of 50 lecturers from AIFCE and 60 lecturers from IMSU Owerri the guidelines they use to

determine the length of time they assign to their students for writing their courses' examinations. 96 (87.3%) of the lecturers recited as it were the regulations of their institutions as regards the number of questions in an essay test and the number of hours to assign to the taking of the test. The summary of what most of the lecturers said is that answering three essay questions should take two hours while writing four questions should be assigned three hours. Eight of the 110 lecturers sampled say that the lecturer who set the question should time himself/herself to take the examination and double the time it takes him/her for the testees/learners. The remaining six lecturers interviewed are of the opinion that nobody has arrived at how to decide the optimal testing time for any test. Deciding how to assign the optimal time to any test is a real challenge.

Inequity involved in asking Testees to Answer r Items, say, out of n items that make up a test/ an examination,

The fact is that when testees are given the option of answering r questions out of n questions, there are n combination r sets of r options from which a testee may choose one. The items of the test are not of equal strength/deficiently. The length of writing to answer each of the questions may not be equal. Thus: assigning the same length of time for writing/responding to each of the unequal combinations of items and assigning the same minimum and maximum scores to the marking/grading of the answers/responses to each of the combinations constitute inequity.

Sitting arrangements during test sessions

When students are allowed to choose those who sit with them during examinations, examination misconducts increase more than otherwise. Those who, did not prepare well for (by some omission/commission should not pass) the examination, may by engaging in some form examination misconduct pass the examination. That is clear inequity.

Sometimes testees are seated very closely to each other, in such a case: (i) they may be inconvenienced and distracted and perform poorly or (ii) they do "combined business" and perform better than otherwise.

Test Items Arrangements/Rearrangement

The ways and manner in which test items may be arranged are many and may also depend on test type. For essay tests, the number of test items in each test is few even when there are sub-items. Each item is presented together with its sub-items. So in an essay test, the items may be presented/arranged in:

- i. an ascending order of item; difficult, complexity or length of time it will take
- ii. a descending order of (as in (i))
- iii. a random order and
- iv. an order representing the themes or units of the curriculum from which the items are drawn.

In a multiple choice test the number of arrangements/rearrangements could be ingeniously very large. Any of the arrangements as presented under essay test

may be rearranged by repositioning the options (or the key).

Test item arrangement or rearrangement though a ploy/an approach to prevent or discourage copying in class tests or public examinations, it poses a challenge in the sense that each rearrangement may be peculiar in the number of stimuli and reactions it present to or elicits from a testee. Thus the various rearrangements may constitute different tests thereby constituting inequity.

Identifying Legitimate Testees

In international examinations/public examinations each having a large number of testees/examinees, examinations for very large classes (coupled with computer frauds) legitimate testees identification is a real challenge to equity. This is the case since someone else could write an examination for another one.

Some Innovative and Suggested Researches for Overcoming the Challenges

A number of innovative research works have been carried out to minimize or overcome some of the challenges outlined above. Also some suggested research topics that could, when carried out, reduce or overcome some of the challenges are briefly presented.

Overcoming The Non-coverage of the Domains of Learning Outcomes in Learner Assessment

A number of research works have been carried out in which projects in various subject areas have been developed.

Some of the projects are used for teaching learning and learner assessment.

When students undertake such projects, the cognitive and psychomotor aspects of learning outcomes are readily scored using the attached scoring guide. The affective aspects of learning outcomes are elicited using an attached affective response rating scale. The rating scale enables the teacher to score the learners' affective learning outcomes wholistically (Ogomaka 1989)

The assessment of learners in the affective domain and character can be carried out using sociometric approaches and anecdotal records. Proper conduct of continuous assessment goes a long way to covering all the domains of learning outcomes in Learner Assessment. Studies should be conducted to find out how school could establish low cost/(improvised) centers for physical fitness exercises. Such centers can improve the learners' physical fitness and help to assess the learners' psychomotor learning outcomes. It is noteworthy that learners should only be assessed in psychomotor areas they are capable in participating.

One other area/issue of challenge is the spread of cognitive test items over the various categories of the cognitive domain. First what proportions of items of a cognitive test should belong to/assess the lower cognitive skills and the higher level cognitive skills. Secondly, what proportion of test items should be appraising the various specific categories. This researcher is not aware or sure of any convention but there are suggestions (as rules of the

Global Learner Assessment Challenges and Some Innovative Researches For Overcoming Them - Professor P.M.C. Ogomaka

thumb) made with regard to the primary, secondary and tertiary levels of the education system. As at the present in Nigeria, levels are further subdivided into the; lower basic, upper basic, senior secondary, lower tertiary and the upper tertiary levels of education. The researcher, in real terms suggests as follows:

Table 1 percentage distribution of items of a test in the cognitive domain to;

Education levels (using Nigeria as example)	Lower level Cognitive Skills (LLCS)		Higher level Cognitive Skills Total (HLCS)				Total
	Kn/ Rem	com/ und	App/ App	An/ Ana	Syn/ Eva	Eva/ Creat	
	()	()	()	()	()	()	
Lower Basic	50%	30%	20%	—	—	—	100%
Upper Basic	30%	40%	20%	10%	—	—	100%
Senior Sec	20%	40%	20%	10%	10%	—	100%
Lower Tert	10%	30%	20%	20%	10%	10%	100%
Upper Tert.	—	20%	20%	20%	20%	20%	100%

Coverage of Curriculum Content

Almost all teachers [especially lecturers] in Nigeria for instance do not cover the various curriculae contents they teach, in assessing their learners/students/pupils. Also external examinations bodies (West African Examination Council (WAEC) and National Examination Council (NECO), as examples, do not cover the various national

curriculae or their various subject syllabuses, (Ogomaka, 1989)

Most of the teachers are neither knowledgeable about the construction and the use of table of specification (Blue print)

Overcoming the Bookish Nature of Learner Assessment in Schools and by External Examination Bodies

The number of items in school examinations and public/external examinations that require examinees/testees to analyze, critique, summarize, evaluate/discuss routine, situations/conditions, definitions or existing passages are too much. What are needed to enable students/learners transfer knowledge, skills and dispositions from schools to real life situations outside school ought to be leading learners to react/interact with real life or simulated problematic life situations/conditions. Some of the times "open book examinations (with timing relaxed)" should be used at the tertiary school level to mirror real life situation and overcome the challenge of the optimal time allowed for an examination.

Furthermore overcoming the challenges of optimal time for learners to write a test; a number of suggestions are given as follow: (i) the teacher/chief examiner should write the test fully including all the expected alternative answers to all the items, the examiner should carefully time himself/ herself; and (ii) assign a maximum of three times; two and half times; two times; one and half times, and equal the length of time it took the examiner to answer the test to learners

at the lower basic, upper basic, senior school, lower tertiary and upper tertiary levels respectively. However the above is a suggestion from this researcher who has taught learners at all the (equivalent) levels as above. It is important that every researcher developing a test to determine empirically the optimal time for taking the test* The determination of the optimal time(s) for the (normed groups of) testees may be guided by the specifications of the Educational Testing Services, that a test is unseeded if virtually all the testees answer at least 75% of the items and 80% of the testees reach the last question [perhaps their respective last questions] (Allen, Brent, Cahalan and Trapani, 2005),

- Giving test item options (as has been explained creates multiple different tests) should be avoided. The time given for taking the test/ examination should be extended appropriately and the testees asked to respond to all the items
- Close or clouded and learners' choice of sitting arrangement during test/ examination sessions lead to perpetration of misconducts/ cheating. Testees should be well spaced out during test. Testees choice of where and who to sit close to should be avoided.

The choice between serialised and randomised sitting arrangements should be determined by the research.

- The use of identity cards/ passport photographs together with finger prints readers should be employed to identify legitimate testees for international and

national test/ examinations. For large classes school/ course examinations ID cards, photo-albums should be used. Also cross checking writings in notebooks and answer scripts should be used when doubts arise or whenever possible.

- The use of items rearrangements in tests reduces/ eliminates copying by testees in examination halls/ rooms but there are still doubts whether the rearrangement brings different tests (Ogomaka, 1999;)

Conclusion and Summary

Learner assessment challenges abound locally, nationally and globally. This lead paper presented some of the global learner assessment challenges and some innovative research works and suggestions to overcoming some of the challenges. Some of these global learner assessment challenges are: the non-comprehensive and incomplete coverage of learner assessment in all the domains of learning outcomes; the non coverage of curriculae contents in learner assessment and the education systems, inability to eliminate some inequities inherent involved in some global learner assessment processes and procedures. Some already conducted innovative research works, suggested innovative research and practical activities to partly or totally overcome the presented challenges are also presented or proposed.

These already conducted or suggested future research works and practical activities are: learners conducted projects for assessing all the learner's leaning outcomes in the various domains; the use of open book examinations; ways of

determining optimal time for writing examinations; ways of identifying legitimate testees; suggestions as to investigate the use of items rearrangement of test items; avoidance of giving test items options; suggestions for providing more avenues for assessing psychomotor outcomes of learners; and the overall curricula objectives. Indeed when the recommendations of the existing innovative research works and practices and the future ones are put in place by all concerned the presented challenges will be overcome

References

- Ainsworth, A. (2015) classical test theory and reliability. www.csun.edu/~ata20315/psu427/topic04-CTF-Reliability.pdf. Retrieved 14/10/15
- Alten, B.M, Brent, B; Cahafan, L, and Trapanic C. (2005) The impact of extended time on SAT test performance. Retrieved from www.collegeboard.com. April 2016.
- Nwogu, B.G. (2016) Averting pedagogical failure in science: insight from educational measurement and evaluation. 103rd Inaugural lecture of the University of Nigeria Nsukka.
- Ogomaka, P.M.C. (1989) projects in junior secondary school mathematics: implication for teaching and continuous assessment, Unpublished ph.d dissertation, faculty of Education University of Nigeria, Nsukka. Nigeria.
- Ogomaka, P.M.C, (1999) strategy toward vision 2010 for preventing malpractices during multiple choice test sessions. Nigerian journal of educational innovations. Vol.1 No. 2.
- Ogomaka, P.M.C. (1996) improving the performances of self-employed artisans craftsmen and technicians (SACT) through itinerant mathematics teaching (IMT). Nigerian journal of educational innovations. Vol1 No. 1.
- Ogomaka, P.M.C. (2006) some grey areas of external examination bodies evaluation of secondary school students achievements. In P.M.C Ogomaka and D.A. Onyemezi (Eds). Book of proceedings of the 20th congress of the Nigerian academy of education Owerri: piece wise.
- Ogomak, P.M.C. (2013) some learner assessment inequalities in Nigerian schools: suggested remedies towards quality assessment. Journal of studies in education vol vii no. 1.
- Ogomaka, P.M.C. and Osuala, R.C (2015) inequities in reporting Nigerian students performances and achievements in the cognitive domain: some remedies. Nigeria journal of educational research and evaluation. Vol. 14 No. 1.

Global Learner Assessment Challenges and Some Innovative Researches For Overcoming Them - Professor P.M.C. Ogomaka

Ogomaka, P.M.C. (2016). Aligning reliability conceptualisation and theories with extant and innovative reliability coefficients conceptualizations. A lead paper presented during the 18th annual conference of ASSEREN held at Imo state University-Owerri, Nigeria 11th to 15th July 2016.

Tracy, B, (2015). Blooms Texonomy of learning Domains, <http://www.nwlink.com-donclark/hrd/bloom.html>.