

POLICY EVALUATION OF ACADEMIC ACHIEVEMENT AND ITS IMPACT ON STUDENTS' LEARNING PROFICIENCY

Abdelaziz Mohammed

English Dept. Faculty of Arts, Social & Management Sciences, Yobe State University,
Nigeria

Abdalazem Osman Gamaraldin

Department of Education Faculty of Arts, Social & Management Sciences, Yobe State
University, Nigeria

ABSTRACT: *This paper aims to investigate the policy of assessing student achievement used in Saudi Arabia which called "continuous assessment/evaluation", (i.e. it is the follow-up and testing the proficiency of the student's skills) as placed by the department of Educational Supervision in Ministry of Education. Moreover, this research paper evaluate the process of students' assessment and its effectiveness to overcome the obstacles and propose appropriate solutions of assessment methods that suit the Saudi students' nature and environment. One main question arises here. It is "Is this programme commensurate with Saudi social and cultural society?" this question is divided into two sub-questions. They are "Does it fulfill the desired standard goals?" Then, "What are the advantages and disadvantages of the assessment used?" Therefore, it is necessary to assume and find answers to these questions. The method used is the historical and descriptive analytical method. The samples participated in this study include 40 teachers and supervisors work in the Dawadmi General Directorate of Education in Saudi Arabia. A 25-item questionnaire is distributed to the sample to sum out the participants' point of views about the goals, advantages, and disadvantages of the assessment method used in Saudi Arabia. The most important result is that this programme needs modifications to suit the Saudi nature and environment. Moreover, finds out that this programme makes students dependent learners. Therefore, the study, mostly, recommends that it needs intensive efforts and experiments to study the Saudi situation in order to find suitable pedagogical methods of assessment that suit Saudi students' nature; and administering intensive training programmes for both teachers and supervisors.*

KEYWORDS: Applied Linguistics, Evaluation Policy

INTRODUCTION

The first author of this paper had been teaching EFL in both school and university levels in Saudi Arabia for more than 10 years before he have joined Yobe State University. No doubt, any teaching process should be followed by an assessment of an academic achievement that lets educational policy makers reach the goals they set. To ensure a better result, plans should be assessing student based on their cultural and intellectual reality of their social surroundings where student resides. The recent assessment approved for judging students' evaluation in Saudi Arabia is called continuous evaluation. This system of evaluation is first announced and approved by Skrevi, 1967, (cited in Abu Libdah, 1982: 128). It is the follow-up and testing the proficiency of the student's skills as placed by the department of

Educational Supervision in Ministry of Education. Teachers follow in accordance to the schedule for the academic year. This programme starts from the first grade primary school and continues until third grade secondary level. Al-Zyoud, (1998, 65) states that continuous assessment was carried out by the teacher during the learning process which begins at the very starting period of learning and going on during the course duration. In the sense that there are no formal tests programmed within the schedule of the academic year, and teachers test skills one by one for each student individually. This sometimes makes many teachers not to follow closely, because it requires effort from both students and teachers.

General Perspective on Examinations

Defining the term test in a dictionary (informally exam) is an assessment intended to measure a test taker's knowledge, skill, aptitude, physical fitness, or classification in many other topics (e.g., beliefs). Bodde (2008) states that a test may be administered verbally, on paper, on a computer, or in a confined area that requires a test taker to physically perform a set of skills. Bodde (2008) adds that tests vary in style, rigor and requirements. So, a test may be constructed formally or informally. Informal test would be a reading test administered by a parent to a child whereas a formal test would be a final examination administered by a teacher in a classroom. Formal testing often results in a grade or a test score. Generally, there are two well-known types of tests. They are standardized test and a non-standardized test (Thissen & Wainer, 2001: 1).

A standardized test is any test that is administered and scored in a consistent manner to ensure legal defensibility. Standardized tests are often used in education, professional certification, psychology, the military, and many other fields Kazin (2010: 9).

Moreover, Kazin (2010: 9) states that a non-standardized test is usually flexible in scope and format, variable in difficulty and significance. Since these tests are usually developed by individual instructors, the format and difficulty of these tests may not be widely adopted or used by other instructors or institutions.

Comparing the two types, standardized tests are widely used, fixed in terms of scope, difficulty and format, and are usually significant in consequences. Standardized tests are usually held on fixed dates that determined. Even though there is little variability between different copies of the same type of standardized test, still there is variability between different types of standardized tests.

In contrast, the frequency and setting of administering non-standardized tests are highly variable and are usually constrained by the duration of the class period. A class instructor may administer a test on a weekly basis or just twice a semester. Depending on the policy of the instructor or institution, the duration of each test itself.

Historical Background

Past Era

Ancient Chinese are the first nation used a nationwide standardized test which was called the "Imperial Examination". This test was established by Sui Dynasty in 605 AD. The main purpose of this examination was to select able candidates for specific governmental positions

Bodde (2005). Later in 1905, this examination was eliminated. In 1806, England had modeled this examination system to select specific candidates for positions in Her Majesty's Civil Service. Bodde (2005) reveals that this examination system was later applied to education and it started to influence other parts of the world, as it became a prominent standard as a regulation to prevent the markers from knowing the identity of candidates, of delivering standardized tests.

Modern Era

During 19th century, universities began to organize written examinations to assess the aptitude of students. From 1842 examination in Cambridge University, leading the professional transitioned to the modern mass-education system. Moreover, the style of examination was fixed with the stress on standardized papers tests to be taken by large numbers of students. In this regard, England in the mid of the 19th century during the development of civil, began to move toward a meritocratic basis by administering standardized tests (Bodde, 2005).

In the European education system, written examinations have been ignored before 1702 (Bodde, (2003: 9). Again, Bodde (2003, 9) ensures that the method of standardized examination began to influence in British universities since 1850s, where oral examination had been the standard. Followed by US, by the educational reformer Horace Mann, carried the shift enormously to helped to move education into the modern era, by modifying curricula in the sciences and humanities, creating a rationalized method for the evaluation of teachers and institutions and creating a basis for the streaming of students according to ability (David,2002: 158-159). In 1952, the first Advanced Placement (AP) test was administered to begin bridging between high schools and colleges.

Some countries such as the United Kingdom and France require all their secondary school students to take a standardized test on individual subjects for the General Certificate of Secondary Education (GCSE) (as in England). These tests are used primarily to assess a student's proficiency in specific subjects. More or less, high school students in other countries like United States, standardized test may not require to be taken to graduate.

Grades or test scores from standardized test, as in the International Qualifications - University of Oxford and Harvard College Admissions, may also be used by universities to determine if a student applicant should be admitted into one of its academic or professional programs. For instance, undergraduate students in the United Kingdom or United States may be required by their respective programs to take a comprehensive examination as a requirement for passing their courses or for graduating from their respective programs.

Finally, standardized tests were found worldwide to be useful to evaluate proficiencies and students' individual differences. Unfortunately, no formal test is administered in Saudi schools levels.

Assessment Methods

Assessments must be designed to reflect the variety of achievement targets that underpin standards: mastery of content knowledge, the ability to use knowledge to reason,

demonstration of performance skills and product development capabilities. There have been encountered thousands of different assessment types. In the light of the variations of assessment, Stiggins et al, (2004) mention only four basic categories of methods:

1. Selected response and short answer.
2. Extended written response.
3. Performance assessment.
4. Personal communication

They are correlated highly with the learning objectives and the intended use of the information.

Selected Response

Selected response and short answer methods are in which students select the correct or best response from a list provided. The formats include multiple choice, true/false, matching, short answer, and fill-in questions. Although short answer and fill-in-the-blank do require students to generate an answer, they call for a very brief answer that is counted right or wrong, so they include these options in the selected response category. For all selected response assessments, students' scores are figured as the number or proportion of questions answered correctly.

Extended Written Response

Extended written response assessment requires students to construct a written answer in response to a question or task rather than to select one from a list. It is at least several sentences in length. It can be one or more such as Comparison, Analyzing, Interpreting, Solving and explaining or Describing in detail.

Correctness of extended written responses can be judged by applying one of two types of predetermined scoring criteria. One type gives points for specific pieces of information that are present. Such as students in an English class are asked to describe directions, points might be awarded for noting that the direction describes the roads, stores, buildings surroundings correctly from the beginning point until reaching the target place. The second type of criteria can take the form of a rule, such as a general rule for making comparisons, which can be applied to any exercise calling for comparison. Scores therefore also take one of two forms: number or percentage of points attained, or rubric scores.

Performance Assessment

Performance assessment is assessment based on observation and judgment; it can be done using: a) Complex performances, such as speaking a foreign language, reading aloud with fluency ...etc. In these situations it is the doing -the process- that is important, b) Creating complex products such as a term paper, a lab report, or a work of art. What counts here is not so much the process of creation (although that may be evaluated, too), but the level of quality of the product itself.

Performance assessments - like extended written response assessments - have two types: a) performance task, b) or exercise and a scoring guide. The scoring guide can award points for specific features of a performance or product that are present, or it can take the form of a heading, in which levels of quality are described.

Personal Communication

It is to find out what students have learned through interacting with them. Such as: looking at and responding to students' comments, asking questions during instruction, interviewing students, listening to students as they participate in class, or giving examinations orally. It looks as informal, rather than formal assessment (in which results are recorded for later use). Often it is, however, as long as the learning target and criteria for judging response quality are clear, information gathered via personal communication can be used to provide descriptive feedback to students, for instructional planning, and for student self-reflection and goal setting. If planned well and recorded systematically, information from personal communication can be used as the basis for assessments of learning. Responses are assessed in one of two ways. a) When questions are asked to require students to provide a simple, short answer, and all what are looking for is whether the answer is correct or incorrect. This is right to scoring for written selected response questions. b) On the other hand, student oral responses are longer and more complex, similar to extended written response questions. So, the evaluation of the quality of oral responses can be done by using documentation or scoring guide. Moreover, complicated responses would occur, for example, during oral examination or oral presentations.

No single assessment method is greater than the others. All are viable options depending on the learning targets to be assessed, the purpose of the assessment, and special student characteristics such as age, language proficiency, or specific learning disabilities. Every assessment development should proceed: 1) identifying the purpose, specifying the targets, selecting appropriate methods, deciding on relative importance of the targets and sampling well; 2) writing the questions using guidelines for quality; 3) eliminating as many potential sources of bias and distortion as possible; 4) administer the assessment; and 5) examining the results for areas needing fine tuning. By doing so, there can be confidence that assessments are being yielding accurate results.

Continuous Evaluation in Saudi Educational Institutions

Policy of assessing students achievement used in Saudi Arabia is called "continuous assessment/evaluation", (i.e. it is the follow-up and testing the proficiency of student's skills continuously) as placed by the department of Educational Supervision in Ministry of Education. Teachers follow in accordance with the schedule for the academic year. This applies began from first grade primary school and continue until third grade secondary level. In the sense that there are no programmed tests schedules during the ongoing academic year. Therefore, teachers test those skills one by one for students individually. Sometimes this makes most teachers weakly uncovered closely, because it requires effort from both students and teachers. Many times, it causes excessive continuous pressure on students' parents because teachers are exaggerating the amount of daily duties (homeworks, assignments) required from the students.

This chronic problem provokes academics in Saudi Arabia to concern with it seriously. So, it was defined by the Ministry of Education as "an ongoing educational process designed to make a judgment of student's academic achievement" (Teacher's Guide in Evaluating the Primary Level Students, 2008: 6). It also called formative or structural evaluation.

Even though, this assessment method was adopted in Saudi Arabia by the year 1998, relatedly, several studies were conducted. Alnajim (2000) conducted a study to judge the feedback of the continuous assessment of the legitimacy science courses which concluded most important results. Significantly, this study concluded that only one goal of the continuous assessment was reached (i.e. it reduce fear and tests anxiety of students). The most commonly used methods were class discussion and observation carried by teachers to assess students. His study also showed that most of the teachers were only concentrated on judging the cognitive dimension of their students. Unexpectedly, the result of this study showed that 90% of the teachers did not get training courses for this assessment method. However, this study was carried out only after one year of applying this method of assessment in schools.

Another study carried out by Beheiri, (2005) to justify the problems face the continuous assessment in high schools. Beheiri, (2005) states that the greatest problem that faces teachers in carrying this method appropriately is the great number of students per class which reduces the accuracy of continuous assessment, taking into account the completion of the curriculum. Again, this study shed light on the improbability of achieving the goals that set by the Ministry of Education is that about 68.8% of teachers did not attend training courses about the continuous assessment. Moreover, Mona Al-Mutairi, (2011) states that critical problem of applying continuous assessment in teaching jurisprudence in primary school girls, is the intensity of topics included in the curriculum, its concentration on the theoretical side only, and the tendency of students on recording information without understanding. So, this case is typically found in English course especially in the college requirement programs (i.e. there are courses that the actual credit hours approved for students to be taken within one term (semester) exceeds 6 hours until reaches 14 hours in others).

A recent study by Sehaim (2011) aims to identify the extent of objectives of the continuous assessment fulfilled in upper primary stage, as well as to judge the positive and negative aspects that resulted after the adoption of its application, and standpoint of teachers and administrators about the difficulties that faced evaluation method.

In fact, Sehaim's study is considered the most important and comprehensive one on this issue. It has been carried out after more than ten years of continuous evaluation. Sehaim's and other studies have shown that there were not enough training programs for teachers concerning continuous assessment. Moreover, there was no plan for any kind of evaluation program as well as continuous assessment. Also, students have not acquainted with this kind of evaluation in the intermediate level. In this study, Sehaim used descriptive method. His sample consisted of (60) supervisor and (364) teachers. He distributed (45) items about the Goals of the Continuous assessment, advantages, disadvantages, and difficulties. The results concluded from his study showed that almost half of the total samples (48%) out of the teachers have training courses in the

continuous assessment representing as well as only 31% of the supervisors have received training courses.

Moreover, the Ministry of Education launched a plan to measure the achievement of students in primary grade. This has been decided after a complaint from the decreasing level of students. There is a tendency in the ministry for the establishment of an independent body to evaluate school levels.

The objectives and the benefits of continuous evaluation

The objectives and the benefits of continuous evaluation (Hussein, 1991: 198; Hameedah and Fatima, 1992: 152) as constructed by the Ministry of Education is that it can:

1. Provide structured and continuing/ feedback for both teacher and student.
2. Show how students master certain skill before moving to other.
3. Stand over the growth of information and knowledge of students and increase them.
4. Identify weaknesses of students that need treatment and foster them in the future.

Methods of continuous assessment were approved to be taken are: quizzes, grading check list, observation, homework and follow-up, classroom exercises, and class discussion through questions and dialogue (Mansour, 1996: 22).

Ministry of Education holds this method and applies it gradually. In 1998, it began with oral courses, issuance rules and instructions for evaluating beginner grades students at the primary level, and then new regulation accompanied by an explanatory note in 2006. Then, this continuous assessment approved to include all grades in primary school as follows: a) applied to the fourth grade in 2006/2007, b) applied to the fifth grade in 2007/2008, and c) applied to the sixth grade in 2008/2009 (Sehaim, 2010).

Tools and Procedures

The historical and descriptive analytical method is used to investigate the goals of continuous evaluation besides the advantages and disadvantages of this programme as well as the effect of the programme on students' achievement. The tool used in this study is questionnaire to investigate the goals, advantages and disadvantages of continuous assessment used in Saudi schools and universities to measure students' achievement. The questionnaire contains two sections. Section one, asks participants about rank and experiences. In section two, there are three dimensions. They include continuous assessment goals, advantages and disadvantages of this programme. 3-Likert point scales are used to get the views of this programme. Yes response was ranged (3) and no coded (1) with neutral scale which coded (2). The duration of this survey is done during the first semester of the academic year 2013-2014. The questionnaire is distributed to 30 teachers and 10 supervisors to sum out their points of view about the continuous assessment and its effect on students' achievement. The tool items show a consistence of reliability correlation of (.75).

Table 1. Items means

Rank	Supervisor		10		Teacher	30	Mean
Experience	More than 10	27	5 to10	8	Less than 5 years	5	
The Goal of the Continuous Assessment <i>is to...</i>							
Make students keep information for a longer period.							2.328
Give an estimate level of students' achievement at the mean time.							2.138
Show families the academic achievement progress of their children.							2.103
Reflect the extent level of skills that student masters and achieves immediately.							2.017
Improve students' assessment procedures.							1.966
Advantages of continuous assessment							
Gives the students a clear picture of the extent of their mastery of the skills.							1.759
Divides students according to individual differences.							1.707
Allows students to apply skills that have been learned previously.							1.690
Develops the students' self-evaluation skills.							1.690
Consider both the hardworking and lazy students as the same.							1.672
Helps the teacher to monitor the progress of his students.							1.621
Allows using varieties of methods of evaluating that fit the educational situation.							1.535
It helps in students' success.							1.517
Eases students' tests fear and anxiety.							1.432
Helps to reduce the problems of students' failure.							1.345
Disadvantages							
Weakness of the academic knowledge of teachers.							1.707
Non- inclusive of assessment skills for less-classroom activities.							1.500
Multitude of fields and skills lead to weakness of teachers to follow their students.							1.483
Lack of interest of both students and parents in the outcome of evaluation.							1.483
Frequent administrative burdens on teachers.							1.448
Weakness knowledge of the assessment methods by teachers.							1.414
Lack of follow-up by school administrators to the application of the methods of continuous assessment in the classroom.							1.328
Increases class-load for each teacher.							1.328
Lack of specialists in the field of Educational Measurement and Evaluation in the departments of education to oversee the evaluation process.							1.328
The large numbers of students in each classroom.							1.310

DISCUSSION AND RESULTS

Concerning the "Goals of the Continuous Assessment" used in Saudi Arabian schools, after calculating the samples' responses (see table 1 above), it is clear that most of the respondents highly accept the goals that set for this programme (over all mean is 2.11) which means that it can fulfil its goals. In putting subjects' responses in descending order, one can see that the item which mentions "*Makes students keep information for a longer period*" (item 5), comes first, which shows a mean (2.32). Then item 3, which states "*Give an estimate level of students' achievement at the mean time.*" the mean is (2.14). Moreover, item four "*Improve*

students' assessment procedures." accepted at the last level but generally, all items were accepted because the statistical results of the responses show a mean preceded (2).

Concerning the advantages of the continuous assessment, it seems that this programme may be useful to apply correctly if teachers spend an excessive effort to test skills separately and put in to their minds there are individual differences need to be aware. However, unexpectedly, respondents are hesitating to give an appropriate yes or no response. Accordingly, teachers and supervisors participated in this study are not agree with this point. The items' calculation that state "*the continuous assessment reduces the problems of student's failure and eases students' tests fear and anxiety, which leads to students' success, shows individual difference and develops self-evaluation skills*" (the overall mean=1.6). Moreover, the study shows that the disadvantages of this programme, all participants are hesitating in responding positively or negatively towards program. From the calculation we can notice that the mean is (1.433), (i.e. this programme needs excessive modification to suit the Saudi context). In conclusion, we can say that this programme is not suitable to give appropriate feedback of students' achievement. Nevertheless, the goal of the continuous assessment is accepted. Finally, the study results support the findings achieved by Alnajim (2000) is that teachers' focus is only on the cognitive dimensions without carrying out the other learning objectives dimensions like Affective and Psychomotor dimensions.

After calculating and analyzing the responses of the participants concerning the goals, advantages and disadvantages, the study finds that this programme more or less cannot serve any academic evaluation in Saudi Arabia unless it followed by formal examinations, and teachers and supervisors well trained to hold this programme. Moreover, the programme puts a strong pressure on both students and their parents because teachers sometimes excessive the assignments that given to students. In addition, it makes students dependent learners. Therefore, the study recommends that this programme, if modify: a) can make students keep information for a longer period; b) give an estimate level of students' achievement at the meantime; c) show families the academic achievement progress of their children; and, d) reflect the level of skills that student masters and achieves immediately. However, unexpectedly, the general not helpful in improving students' assessment procedures.

REFERENCES

- Abu Libdah, Alsaba' Muhammad. (1082). Principles of psychometrics and educational assessment. Oman: collaborative presses.
- Alnajim, Mohammed. (2000). The Reality of Continuous Assessment in the Teaching of Forensic Sciences: Teachers and Supervisors in Primary School for Boys in Riyadh City. Unpublished Master's Thesis. College of Education, King Saud University: Riyadh
- Al-Zyoud, Nader (1998). Principles of Measurement and Evaluation in Education. Amman: Dar Alfikr.
- Beheiri, Mohamed Hamed. (2005). Continuous calendar application problems in teaching the Koran high school. Unpublished Master's Thesis. College of Education, King Saud University: Riyadh
- Bodde, Derke (2005). "China: A Teaching Workbook". Columbia University.

- Bodde, Derke., (2005). Chinese Ideas in the West. A teaching Workbook". Columbia University.
- David R. Russell (2002). Writing in the Academic Disciplines: A Curricular History. SIU Press. pp. 158–159.
- Hameedah and Fatima. (1992). Mastery Learning and its Effect on Academic Achievement. Journal of Educational Studies folder (7). Riyadh.
- http://www.collegeboard.com/prod_downloads/about/news_info/ap/ap_history_english.pdf" GCSEs
- Hussein, (1991). Formative assessment in teaching and its effect on academic achievement. Journal of Education, Qatar National Committee for Education and Cultural Rights. 243-265.
- Kazin, Edwards, Rothman & Walker, David (2010), 142. (2003-07-09). "Fair game". London: The Guardian. Retrieved 2003-07-09. Bodde, D., Chinese Ideas in the West, p.9
- Mansour, Abdul Majid. (1996). Educational Evaluation (Basis - Applications). Jeddah: Dar Tihama..
- Mutairi, Mona. (2011). Most Important Problems of the Application of Continuous Assessment in Teaching Jurisprudence in Primary Education for Girls: Teachers and Supervisors in Riyadh. Unpublished Master's Thesis. Faculty of Social Sciences, University of Imam Muhammad bin Saud: Riyadh.
- Saudi Ministry of Education (2008) Teacher's Guide in Evaluating the Primary Level Students. Saudi Arabia. Riyadh.
- Sehaimi, Turki (2011). The reality of continuous assessment for students in the upper grades primary in forensic science courses. Unpublished Master's Thesis. College of Education, King Saud University: Riyadh.
- Stiggins, R., Alter, J., Chappuis, Jan and Chappuis, Steve (2004). Classroom Assessment for Student Learning: Doing it Right-using it Well. Part II. Assessment Training Institute. Retrieved from: www.assessmentinst.com
- Thissen, D., & Wainer, H. (2001). Test Scoring. Mahwah, NJ: Erlbaum. Page 1, sentence 1.

APPENDIX

Questionnaire

Dear colleagues, you are kindly required to respond according to your own point of views about the continuous evaluation used in Saudi educational institutes (at both university and school levels). The questionnaire contains two sections: **Section one**, includes information about the participants such as rank and experiences. **Section two**, is divided into 3 groups of

Rank	1-Teacher	2- supervisor.....	
Experience.	Less than 5 years.....	5 to10.....	More than 10.....

questions. They measure goals, advantages and disadvantages of the continuous evaluation.

Section 1: Tick what suits your rank and experience in the table below:

Section 2: Write 1, 2, or 3 in the box in front of each item as explained in the boxes below:

3 = Yes	2 = Not sure	1 = No
----------------	---------------------	---------------

The goals of continuous assessment are to		
1	Show families the academic achievement progress of their children.	
2	Reflect the extent of skills at which students master and achievement immediately.	
3	Give an estimate level of students' achievement at the mean time.	
4	Improve students' assessment procedures.	
5	Make students keep information for a longer period.	
Advantages of continuous assessment		
1	Helps to reduce the problems of students' failure.	
2	Eases students' tests fear and anxiety.	
3	It helps in students' success.	
4	Divides students according to individual differences.	
5	Develops the students' self-evaluation skills.	
6	Allows using varieties of methods of evaluating that fit the educational situation.	
7	Gives the students a clear picture of the extent of their mastery of the skills.	
8	Allows students to apply skills that have been learned previously.	
9	Helps the teacher to monitor the progress of his students.	
10	Consider both the hardworking and lazy students as the same.	
Disadvantages		
1	Multitude of fields skills lead to weakness of teachers to follow their students.	
2	Lack of interest of both students and parents in the outcome of evaluation.	
3	Non- inclusive of assessment skills for less-classroom activities.	
4	Lack of follow-up by school administrators to the application of the methods of continuous assessment in the classroom.	

5	The large numbers of students in each classroom.	
6	Increases class-load for each teacher.	
7	Lack of specialists in the field of Educational Measurement and Evaluation in the departments of education to oversee the evaluation process.	
8	Frequent administrative burdens on teachers.	
9	Weakness of the academic knowledge of teachers.	
10	Weakness of the assessment methods by teachers.	