

## **Standardization of the Saudi Secondary School Certificate Examinations and their anticipated impact on Foreign Language Education**

**Dr. Ahmed Gumaa Siddiek**

Dawadami Community College, Dawadami - P.O. Box 18

Shaqra University – KSA 11911

E-mail: aahmedgumaa@yahoo.com, Phone : (+966 55 94 55 198)

### **Abstract**

*Scholastic Achievement tests are standardized in form and content to **fairly** serve all the audience - who take them - under similar circumstances. Therefore, they are reliable tools for measuring the performance of the examinees. On the other hand, proficiency tests are intended to measure the overall abilities of the learners. So they cannot serve as yardstick to tell us about the degree of the attainment of the national educational objectives embedded in the textbooks. The Education authority is advised to recall the national certificate examinations which qualify learners for higher education, under fair national competition. The role of other proficiency tests should be integrative and complimentary to the major role of the Ministry of Education, who alone should have the right to interpret the usefulness and credibility of secondary school certificate, and how to use this national document in the assessment of students' performance at the personal, national or international level.*

**Keywords:** standardization. School leavers, achievement tests, proficiency tests, credibility, assessment, foreign language education

### **1:0 Introduction**

Standardized tests are high-stake examinations used to measure participant abilities and skills. They are so named because their administration, format, content, language, and scoring procedures are the same for all participants, where these features have been 'standardized'. Standardized tests according to Anderson:(1975:384), attempt to provide for measurement of individual differences in as unambiguous ways as possible. Thus, the process of standardization permeates all aspects of testing: construction, administration, scoring, reporting, and interpretation of results. As the Ministry of Education in Saudi Arabia abolished down the General School Certificate Examinations, the Saudi secondary school leavers do not sit for one national standardized achievement test at the end of their secondary school education. However, they have local teacher-made examinations in all school disciplines. They do not go into one integrated and formatted language exam at national level, all over the country. The result of this local testing process has yielded a tangible deterioration in students' English language performance and (lazy) teaching tendency in the teacher's community; as teachers exert little effort on teaching. However, they pay much focusing on teaching test techniques as an end in itself, to enable their clients (students) collect high marks at the end of their secondary schooling, with very little effort or with no effort at all.

These collected marks are unfortunately not all *considered* for university admission, as the requirements to university in Saudi Arabia is now based on the results of the Proficiency Achievement Test (GAT) and Standardized Achievement Aptitude Test (SAAT) of The Notational Testing Centre (Qiyas). Universities consider only 30% of the school grades that the students get at secondary school examinations – where practically learners spend 12 years of general education - and 70% from the (GAT) & (SAAT). These two exams created an environment of dissatisfaction among *almost all* stake-holder from learners, teachers to parents. The Qiyas provides a proficiency test rather than an achievement scholastic test, which is supposed to be absolutely based on the efforts exerted by the learners in studying their school textbooks. So the Qiyas language test is a proficiency test, as it tests the overall language elements, skills and abilities, while it is expected to provide an achievement test based on the syllabus secured by the Ministry of Education at the secondary school stage. Therefore, the Qiyas test lacks content validity which is a major feature in standardized scholastic achievement language tests.

### **2:0 Scholastic achievement tests and content validity**

Teachers can teach their courses with all their might. They might address the needs of their students - and we must be thankful for this job. But when it comes to testing their students; teachers face many unprofessional-intentionally or unintentionally - unethical practices that appear to them. Intentionally or unintentionally, we sometimes do great harm to our students, when we expose them to a testing situation for which they are not prepared or when we give them an exam, which is not well designed and not good enough to serve and secure the purpose of the course- which we have taught them or the unit, which we have just finished. Inaccurate tests are very harmful, to both teachers and students, as well as to the community.

An inaccurate test makes great mess in the whole educational process as it yields negative backwash. Backwash is defined as the effect of test on teaching and learning. The backwash can be positive or negative. See Hughes (1995) Sax (1980), Priest (1996), McNamara (2001). An inaccurate test spoils the entire educational efforts of all the stake holders: learners, teachers, parents, educators and decision makers. It does all of that because it does not help to fairly *judge* the extent of the degree of the attainment of our local, personal or national objectives. But where does the inaccuracy of test come from? Hughes (1995:3) points to two main causes of inaccuracy. The first cause: is the test contents and test techniques and the second source of inaccuracy is the lack of reliability, which is the ability to measure consistently. The technique of the test is sometimes misleading, if the wrong or the unsuitable technique is used to examine one specific item, such as testing composition by using multiple choice test as Hughes (199:3) pointed to that. A test should also have beneficial backwash on teaching and learning as well. If a test is regarded as important then preparations for it can come to dominate all the teaching and the learning activities. And if the test content and testing techniques are at variance with the objectives of the course, then there is likely to be harmful backwash, Hughes (1995).

Standardized achievement examinations are expected to show the degree of students' achievements in the specific course that they had studied and not their language **overall proficiency skills**. This point can be supported by Menke (1998) when he pointed out that, in order for assessments to be effective and useful for educators in instructional practice, they must be deeply entwined with the classroom teaching and learning driven by the standards. It is also supported by Bronwyn (2002) who says, "If tests are aligned with standards and curricula, students will have an increased chance of demonstrating what they know and are able to do." Teachers of ELLs (English Language Learners) need to be involved in the decision-making process regarding which tests will be used. So an accurate test is a good test but the question there still: what is a good test? And what is the best test that can serve and secure our educational goals as language teacher? This question has two parts. The first is what is a good test? Most of language testers have come to some sort of agreement that a good test generally have the following features:

- Reliability
- Validity
- Comprehensiveness
- Practicality
- Positive backwash.

In the Next pages we will try to examine each concept in brief, because this will help in understanding the mission assigned to test and testers as well. The second part of the question is what is the best language exam that can serve our objectives of language teaching? If we believe in relativity, then we can definitely come to answer this question, that there is no absolute best test, but there are suitable test which can serve the task. Hughes (1995), comments to this by saying that, "Language testers are often asked to say what is 'the best test' or 'the best testing technique? Such questions reveal a misunderstanding of what is involved in the practice of language testing. In fact there is no best test or best technique. A test which proves ideal for one purpose may be quite useless for another. A technique which may work very well in one situation can be entirely inappropriate in another. So the question is: what determines the best test or the best technique? The *target* behaviour will determine the test or the technique, which we will use to elicit the desired results from the student." (ibid)

### 2:1 Reliability

Understanding reliability as Sax (1980:255) says is very important for two reasons: First, principles of test constructions depend upon having a clear understanding of this term. A test of low reliability is a waste of time for both teachers and students since it permits no conclusion to be drawn. Second, the selection of test depends, in part, on a consideration of the reliability of measurements. Unreliable tests are no better than assigning students random scores. So a good test is a reliable test which is as Sax (1980:257) defined is the extent to which measurement can be deepened on to provide consistent, unambiguous information. Measurements are reliable if they reflect "true" rather than chance aspects of the trait or the ability measured. A good test is a reliable test; a test that measures consistently. On a reliable test, you can be confident that someone will get more or less the same score, whether they happen to take it on one particular day or on the next; whereas on an unreliable test the score is quite likely to be different, depending on the day on which it is taken. Hughes:(1995:22), sees the reliability of test as the extent to which it is free from random measurement error. Tests that are highly reliable are consistent and trustworthy. The better tests have a reliability of close to 1.00, while very poor tests have a reliability of close to 0.00. Another definition is for reliability is as seen as referring to the consistency of test scores. Various methods are used to calculate and estimate reliability depending on the purpose for which the test is used.

**2:2 Validity**

Technically validity can be calculated according to Sax (1980:258) as true variance divided by obtained variance. In practice, true variance, of course, have to be estimated since it cannot be computed directly. We can have some more definitions of validity. "A test is valid when it measures what it claims to measure." Garrett (1964:30). "A test is valid when it "measures what it ought to measure." Ebel:( 1972:436). According to Lado:(1975:30), validity is the answer to this the following question, "Does the test measure what it is intended to measure? If it does, it is a valid test." Abbott:( 1992:178), holds the view that "A test cannot be a good test unless it is valid. The essence of validity means the accuracy with which a set of test scores the test measures what it claims to measure." At last Hughes:( 1995:22) believes that "A test is said to be valid if it measures accurately what it is intended to measure." So a test is a tool, but this too should be valid and reliable to carry out the job professionally. A valid and reliable tool is the one on which we can absolutely depend on to give consistent measure. A ruler is a reliable tool but it is *impractical* to use in measuring the distance from Riyadh to Makkah, although it can give the same reading all the time it is used but in this context it is not a valid tool.

**2:3 Content Validity**

According to Hughes :(1995:27), a test is said to have content validity if its content constitutes a representative sample of the language skills, structures, etc. with which it is meant to be concerned. Tyler in Ross (1963:114) regards a valid test as the one which affords satisfactory evidence of the degree to which the students are actually reaching the desired objectives of teaching, these objectives being specifically stated as in terms of the kind of behavior. Content validity is an important notion for the measures of skills and knowledge that are frequently employed in evaluation studies. In this context, content validity usually refers to the degree to which a measure captures the program objective (or objectives.). Anderson :( 1975:460) defined content validity by saying that; validity refers to the extent to which the test we're using actually measures the characteristics or dimension we intend to measure. Chair :( 2003), thinks that content validity refers to the extent to which the test questions represent the skills in the specified subject area. Content validity is often evaluated by examining the plan and procedures used in test construction. Validity is the accumulation of evidence to support a specific interpretation of the test results. If the test lacks content validity then it will be harmful to both learners and teacher; because teachers will not be able to secure the specific educational objectives in the textbooks, and learners would hunt for marks instead of doing real learning by studying their textbooks. In this case we would expect a negative washback of tests on teaching and learning.

**2:4 Face Validity**

The term concerns the extent to which the test appears to look like a test of the concept it is intended to measure. The concept is developed by Hughes: (1995:22) as saying that a test is said to have face validity if it looks as if it measures what it is supposed to measure. For example a test which pretended to measure pronunciation ability but which did not require the candidate to speak, might be thought to lack face validity. Face validity is not necessarily equivalent to content validity. Thus a test item may have content validity as judged by an expert in the field but may not have face validity for test users and test takers.

**2:5 Specifications of Educational Objectives**

Assessment cannot be achieved unless we have clearly stated the educational objectives that the student must reach. Goals and objectives must be written appropriately to ensure that they are valuable and feasible. It is important to interpret these objectives into concrete behaviour, which can be seen and observed. But this process may face some problems, Tyler in Ross:(1963.114) summarizes this by saying all methods of measuring human behavior involves four technical problems: (1) defining the behavior to be evaluated, (2) selecting the test situation, or determine the situation in which the behavior is expressed, (3) developing a record of the behavior, and (4) evaluating the behavior recorded. But how can we avoid such problems and construct reliable and valid tests and avoid the inaccuracy in our tests? As we see there are two components of test reliability: the first component is the performance of the students in their different examinations and the second component is the way teachers mark or rate the examinees performance. Hughes suggests some ways of achieving consistent (reliable) performances from the candidates. See Hughes (1995:33)

**2:6 The Educational and Socio-economic Consequences of High-Stake examinations**

The term high-stake means the consequences associated with test results. American families make important decisions, such as where to live, based on the scores from these tests. This occurs because real estate agents use school test scores to rate neighborhood quality and this affects property values. Test scores have been shown to affect housing prices, resulting in a difference of about \$9,000 between homes in grade "A" or grade "B" neighborhoods. At the national and state levels, test scores are now commonly used to evaluate programs and allocate educational resources. Millions of dollars now hinge on the tested performance of students in educational and social programs, (ericass:2003). The results of tests can also be used that way in this country as families move from village to city with their children to admit a university or higher educational institute.

These examinations are benchmarks on the national level, from which many inferences can be drawn whether in the field of education, sociology, economy or politics.

### **2:6:1 Tests in educational context**

The ministry of education or any other educational affiliation will be able to interpret the results of these examinations and see if the educational aims have been achieved as per the results of the students' performance at the national level. So through these examinations the ministry can serve the pedagogical and social philosophy of the country or the government in education. It will also be able to judge the degree of students' mental and physical aptitude to use the information they got from their textbooks, in real life situations. Through these examinations, the foreign language teacher can secure the minimum educational objectives of language teaching to young people, at this specific time of their ages. Also teachers of other disciplines can be sure about how these school leavers would make the practical use of their knowledge in their life. So we will be able to see how the boy or the girl can use mathematics in solving problems of real nature as finding the area, measuring the distance, and using geometry and algebra in solving problems. The science teachers will also be sure that the scientific knowledge will be of practical use for the learner in his/her real life. The same is applicable to humanities and arts.

### **2:6:2 Test in social context**

Achievement high-stake examinations results differentiate schools, institutions and regions from one another. The individual student's performance in examination is considered in most societies. High achievers are highly esteemed for their achievements not for themselves, but his achievement might be considered as victory for the family, if not a victory for the whole tribe. So examinations have their social impact on individual persons as well as communities. They give high prestige to the persons as well as communities.

The Candidates' performance in national examinations is inherent in the rating and status accorded a school, tertiary institution or state due to 'variation in "standards" among schools and institutions. (Bande 2010), for practical reasons believes that, without national examinations, it would be very difficult to assess individuals from different schools and regions and the schools and regions themselves. Although no qualifying examination is based on achievement in the English language alone, English is the gatekeeper or determinant of admission into higher levels, educational, social or political – without an adequate performance in the English language exam an individual is pegged (down) according to Bande, (2001) in Nigeria but this may also be applicable to other people in other places in the third world where the English language is the passport to success and social prestige.

### **2:6:3 Test in economic context**

Human resources are the most important capital in these competitive and challenging modern times. Information is crucial, to help strategists and decision makers to take the right decision. Languages are the means through which knowledge is imparted. Thus language education is a business itself as we can see in IELTS and TOFEL examinations all over the world. Languages as well, are means of conveying knowledge from one society to another, to help promote the human resources, and pave for mutual exchange of commodities, services feelings, ideas and attitudes. According to Bande (2001), a wider impact of poor language teaching and assessment in society is the waste of time and other resources expended on educating people. Moreover, additional cost is incurred by universities and corporate bodies in conducting English proficiency tests, to assess what has already been supposedly tested –certificates are seen as no longer reflecting proficiency or competence levels. Through tests we can judge the qualifications of individual persons and then we can decide to offer him or her a job according to the test result which can show the candidate's ability to carry out a certain job. Therefore, examinations through examinations we can determine the degree of individual fortune in getting well paid job according to his/her qualifications rather than his political, social or ethnic affiliation.

### **2:6:4: Test in political context**

Politically, the ruling regime or government has the right to interpret the results of examinations as part of its political orientation. The government or the state would want to be sure about the attainment of the general and specific educational objectives, which are set as part of the educational strategy of the government or state. Examinations are benchmark of the success or the failure of the educational objectives and thus they may serve as indicators for success of the political orientation. But language examinations can be misused as tools of social exclusion when intentionally make the examinees fail their exams, to serve specific hidden political and cultural agendas. See McNamara: (2001:), (Hughes:1995), Siddiek (2010).

### **2:7 The Saudi School Certificate Examination**

The (SSC) in some respect is similar to those high-stake examinations in the USA, European and Asian countries. It arouses national as well as international interest. At the national level it qualifies students to admission to university as it is part of the requirements to university admission. At the international level it is considered as a yardstick of the student achievement at the secondary school level.

Accordingly, this Saudi School Certificate - among other-criteria - is a valuable document for the measurement of the student's performance in his/her general knowledge in the general education in secondary school. This certificate can serve as a reliable document if new procedures are taken into consideration. This is what I intend to discuss in the next pages.

### **3:0 Standardized Tests and Evaluation**

Standardized tests can be used to measure participant aptitudes, abilities and skills. They are so named because their administration, format, content, language, and scoring procedures are the same for all participants as these features have been 'standardized'. Standardized tests according to Anderson:(1975:384), attempt to provide for measurement of individual differences in as unambiguous ways as possible. Thus, the process of standardization permeates all aspects of testing: construction, administration, scoring, reporting, and interpretation of results. These examinations are sometimes locally developed e.g.(SSC); and they are sometimes commercially available created for most achievement areas and for some aspects of language proficiency such as the international American test of English known all over the world as TOEFL, and the British English IELTS. When considering the definition of "standardized test," it is clear that all high-stakes tests (should be standardized. These tests are useful when selecting people for a particular program because; they are designed to differentiate among the test takers. In addition, norm referenced tests can provide general information that will help us to match classrooms for overall achievement levels before putting them to a particular program. The term standardized test originally meant, and still means when used precisely, a test that:

- Has been carefully, expertly constructed with analysis and revision.
- Has explicit instructions for uniform (standard) administration; and has table of norms (standard) for score interpretation derived from administration of the test to a defined sample of students.
- And "loosely, the term can refer to almost any published test or inventory, whether standardized in the manner just describe or not." Ebel:(1972:465)

### **There are three categories of standardized tests**

#### **3:1:1 Proficiency Tests**

They are designed to measure the peoples' ability in a language *regardless of any prior training* they may have had in that language. The content of a proficiency test therefore, is not based on the content or the objectives of language courses which people taking the test may have followed.

#### **3:1:2 Achievement Tests**

These tests are directly related to language courses. Their purposes are to establish how successful individual students, group of students, or the courses themselves are. Scholastic tests which students take at the end of school year represent the greatest part of these tests. These tests are designed to help in placement of students in school classes or colleges and universities such as this (SSC) examinations.

#### **3:1:3 Diagnostic Tests**

As the name suggests, it is the test that is used to identify the weaknesses and strengths of a students in certain language elements such sounds, intonation, stresses, morphemes and skills such as writing and reading. There are other types of Tests: such as the placement tests which are intended to give information, to help to *place* student at the stage or in the part of a certain teaching programme that is most appropriate to his abilities and the promotion test that may be used in educational experiment. Hughes:(1995) thinks, if we want to compare between two methods and determine which one will fit to lead to better result, a pre-test is given to the controlled group and the experimental group. And the last one is the remedy test which is used to help select those students who need special help and treatment so as to do remedial work to treat their weakness at certain area of knowledge. So we do not need to say that these functions of tests apply to all subjects including the foreign language. According to Al-Khuli :( 1995:95), the same test may perform several functions simultaneously. A test that measures student's achievement may be used by the teacher in self-evaluation as well. It may also be used as a criterion to provide patterns with information about the syllabus success.

### **4:00 The Role of National Testing Center (Qiyas) in Saudi Arabia**

The role of Qiyas is expected to be a helping role. It is supposed to be integrative with the role of the Ministry of Education in Saudi Arabia. But it seems that Qiyas has taken the lead by stealing all the light from the Ministry of Education and cancelling its role as the main provider of educational evaluation in the kingdom, as well as the sole body which is responsible for the educational evaluation of students in this country. The ministry of education provides the syllabus and secures the suitable environment for the educational process to take place in the school. The schools themselves are well equipped to receive the clients an accommodate the for 12 successive years of education. The budget of the general education is very generous as the kingdom allots about 26% of its national budget to education. Teachers are recruited from inside and outside the country to perform teaching.

Teaching is carried out professionally as we have experienced by ourselves as school teachers and supervisors in the general education. But we still complain from the poor outcome of the general education products as resembled by school leavers' performance in practical life. I am not here to recall the reasons behind this deterioration but I will add one new reason to that, it is the new proficiency and achievement test carried out by Qiyas.

**The National Center for Assessment of Higher Education** (Qiyas) proficiency and achievement test Examinations. Qiyas provides GAT and SAAT and special English language test called STEP:

**4:1 What is STEP?** STEP is an acronym for Standardized Test of English Proficiency. It is designed to be an objective and unbiased test of a person's level of proficiency in the use of English. The test is made up of the following four components: Reading Comprehension Structure, Listening Comprehension and Compositional Analysis. According to ([www.Qiyas](http://www.Qiyas)), it is based on the growing international needs for the English language; several academic and non-academic institutions have approached the National Center for Assessment in higher education calling for the development of an English test that could measure the proficiency of their applicants. So it is obvious that the need for this centre comes from the call from higher educational institutes.

This assertions seems strange as these institutes are highly qualified to run such tests within their vicinity, and with their abundance of qualified academic personnel, who are specialists in nuclear physics to psychological measurements. Qiyas proficiency test might give proper measurements to the language standard of these students, but this proved not to help in language acquisition. The proof is that these newcomers whom, the Qiyas certified as got high rates in comparison with their colleagues; do practically need further training in English in the preparatory college, which is designed for this purpose. The Qiyas test in fact does not give any practical information about the candidate linguistic abilities; it gives an overall assessment of all language skills of the examinees. Still the problem is not with the measurement, but with the techniques which are to be followed by the examinees to prepare to sit for these Gat and SAAT examinations. These techniques control and affect the whole process of language education; as *all* the efforts of teachers and students will be focused on learning testing techniques rather than doing real teaching and learning from their textbooks. This is what we can as on the next pages.

#### **4:2 What is wrong with this proficiency test of Qiyas?**

The Qiyas role should be integrative and consultative and not that leading in the evaluation process of school leavers. As we may know standardized proficiency tests presuppose the acquisition of certain *techniques* to do the test. Therefore, most of the time, the learner is consumed up in learning the techniques of answering the items of the test rather than doing real learning. Even the tutorials, on the website focus on giving training to the examinees on how to guess the questions, rather than telling them how to acquire certain learning strategies. So, the Qiyas test puts the student and the teacher on the wrong track. Teachers at schools focus all their might and efforts on teaching the students how to guess the right option. Students focus their effort on learning hunting for the right option; thus both the teacher and the learner totally forgot about focusing on seeking learning from the textbooks, which is supposed to be embedding the syllabus of the ministry of education, and which is expected to secure the minimum knowledge for the student at this time, in such a dose. Technically, a language proficiency test does not presuppose mastering specific details of a certain syllabus or specific content, because it tests the **overall** abilities of the learner's language skills.

Therefore, this Qiyas proficiency test neither tests competence nor tells us about the degree of achievement in the course, which the learner spent 12 years to achieve in the general education. It does not tell us what the examinee can do with his language competence and how he could use this language competence in his further education. Most of my students whom I have unofficially interviewed were not happy with the language examinations of Qiyas. They think this Qiyas test has robbed them 12 years of their life, as they were doing their best to pass their school subjects through an achievement test – based on their textbooks - and not through a proficiency test which may be suitable for jobs selection but not university admission. The universities require 70% of Qiyas Test and 30% from the general certificate achievement test. Of course this is absolutely unethical, uneducational and uneconomical. It is unethical because it has become a **must** for any student to register for this exam as it is considered a part of the admission requirement to university. It is uneducational because the practice for the test dominates the entire process of teaching and learning. Both teacher and learner focus on the techniques of test rather than doing teaching and learning. This practice has totally cancelled the educational and administrative role of the ministry of education, where 26% of the national budget goes to education but the result is to be dominated by Qiyas which was supposed be a unit affiliated to ministry of education and not a rival. It is uneconomical because the entire education budget is allotted to secure only 30% of the educational goals. Millions of riyals are spent to gain nothing as well as other millions paid by the examinees to sit for these exams more than once. This is also unethical.

In addition, time-wise, moneywise and physically; these tests take much of the examinees as some people sit for this exam for more than once. It is mentally and physically exhausting. Parents and many educationists are not happy with the results of this experience. There some students who got more than 90% in their achievement tests but they got round 60% in the Qiyas test. This will need much explanation by both the Qiyas and the ministry of education, because *one* of them must be *wrong* about the measurement of this examinees' achievement. Most educators are not happy with what is happening with Qiyas. Here is a long quotation or actually a translation of one prominent educational figure Dr. A/Aiz Al-Garr-Allah, He says, "the decentralization of secondary school examinations by the ministry of education was a historical decision which aimed to improve examinations process, reform the syllabus, save money and minimize the human efforts which had all been used for the sake of 30%, which the student would get in the second final term exam. The decision was definitely for the benefit of the student because it saved him the burden of centralized secondary school examinations. *But* students found themselves facing *continuous* work for the year's work and the monthly tests, so as to accumulate their marks for the second term.

In addition to that they found themselves faced by the Qiyas examinations where they would sit for four additional examinations. Dr. Garr-Allah continued to say that, and then after all of this [mess] students should have to prepare themselves for the preparatory year where they would run another experience of testing for the college admission. In this way the student will be under the mercy of the ministry of education, the Qiyas and the preparatory college. This is definitely a big burden of examinations which will require time and money as well. It makes these young men exert physical efforts and experience high psychological stress and pressure. Parents complain of the complexity which their children get into to run these tests. Most of the parents complain that the idea of the test presupposes a sophisticated family environment in which parents are literate to handle the new approaches of testing, adopted by the Qiyas. Some parents say that they are unable to help their children to pass these exams because they do not understand these new testing techniques. Even highly educated persons express their inability to help their children, because they were educated in the old traditional way which as they say was not that bad.

Accordingly to Al-Hussein (2010) that a number of teachers in Dammam who were nominated to sit for teaching posts examinations in the ministry of education, expressed their anger from the difficulty of the test. They said the questions were not clear, illegible and indecipherable. They requested the Qiyas to rethink of these tests. Some of them said all their dreams have been deferred to get a job in the public sector. They said they would seek refuge in the private sector. The Director of the centre said that there were 34000 teachers sat for the exam in 19 centres all over the Kingdom. It is not my intention in this paper to give critical academic assessment view point of the work of Qiyas centre, but I would only want to shed light on some aspect which had not been considered. It is clear that the Qiyas centre has totally cancelled the role of the ministry of education as the sole major provider for education to young learners. The admission to university requires 70% of Qiyas and 30% from the general certificate of education in the ministry of education. This is of course illogical it spoils the whole education process.

#### **4:3 Integrative Role**

The role of Qiyas should be integrating partner among other roles of the ministry of education and universities. The test of Qiyas permeates all the educational process. Teachers now confine their job to teach the technique of responding to Qiyas tests instead of focusing on real teaching of the syllabus in the textbooks. Such highly and sophisticated approach of measurement presupposes two things:

1. A highly literate community. This means that the targeted community should be alert and ready to accept the features and results of the new innovation. Qiyas in some way is similar to the American SAAT, GAT and Aptitude Test in Europe and Asia and some other parts in the first world. But the case is different there, because the community is ready to handle this new way of testing.

But the most important point is that all the stake holders are sure that the future candidate of this kind of testing has already fulfilled the minimum requirements of the test at the school, and thoroughly covered the gap of general knowledge. I mean students come to sit for SAAT or GAT Tests already mastered the school subjects, which were supposed to be the minimum requirement of knowledge, that qualify them to sit for such proficiency or achievement tests. Our school performance here is not that matching as we all complain of the poor performance of these schools leavers when they come to university. Therefore the reform begins in the school to do real teaching and learning, and then the test comes next. "The test", as Davies (1968:5) in Hughes (1995:2) believes, is "there to serve the teaching and not to precede its entire activities. The good test is an obedient servant since it follows and apes the teaching." We may not agree with Davies but can support Hughes (1995:2) who sees "the proper relationship between teaching and testing is surely that of partnership." This is exactly what is expected from Qiyas and not to have 70% dominance on the evaluation process of education in the country.

2. The second point is the nature of the client who seeks this service; we mean the student and the university or the education institute. I think both the student and university do not lose much if they ignore this. In addition to that I think the student is not ready to receive such an educational service because we have not prepared him at school to run such an experience. We have to promote the approaches and techniques of testing in our schools first, and then we can expect the student to accept highly advanced batteries of achievement and proficiency measurement mechanism. Actually we have to improve the teaching performance then the testing would come next. The problem is not with the assessment of the learner's performance; the problem is with the basic knowledge given to cover this learner's practical needs. Learners exert big efforts to internalize the approaches of Qiyas testing, rather than doing real learning by studying their textbooks which are supposed to be the minimum dose, that we expect them to assimilate at this age.

### 5:0 Conclusion

Education in the Kingdom needs revision in many domains. First: the objectives of education must be clear and well described and embedded in the syllabus. Then the mechanism of carrying out these objectives should be clearly set. The teachers' selection and training, revision of the syllabus and other educational aspects need serious exploration. After this is achieved we can then come to the mechanism of assessment and evaluation. I think since the ministry of education is the major education provider in the Kingdom, it does not lack the mechanism to carry out national examinations all over the kingdom, to measure the achievement of students, rather than measuring their proficiency, with fair and formatted national examinations in all school disciplines. The solution remains in reconsidering and recalling the General Certificate National Examination. This examination is a national standardized achievement test which will be fair to all students, as standardized tests they can be used to measure participant abilities. The process of standardization permeates all aspects of testing: construction, administration, scoring, reporting, and interpretation of results. These examinations should be under full supervision of the ministry of education, from the moment of the writing, the scoring of the test, up to the moment of releasing the test results to the public.

The ministry of education should be the *sole* body to interpret the results of these tests, and should be the sole body to say how the results of these students' are to be interpreted and used locally or abroad. The right of interpretation of test results should not be given to any person, but it should be the right and duty of the ministry of education in the kingdom. What we practically need is educational reform in syllabus, teacher training and then improvement of assessment mechanism. The Ministry of Education should reconsider to recall the National Standardized Achievement Test once again, because through this test *fairness* can be guaranteed to all learners, as all the students in the country would sit for one *formatted test under the same conditions and within the same educational environment*. Standardized achievement tests can help the ministry of education to attain and secure the educational objectives, as all teachers and learners will all to focus on the syllabus and assimilate its constituents. Then the ministry of education will be able to *decide the usefulness, trustworthiness and credibility* of these test results for the private or public use in the Kingdom or abroad.

### References

- Abbott, Gerry: (1992). The Teaching of English as an International Language. Tomas Nelson & Son. USA
- Al-Khuli, Ali Mohammad: (1999). Teaching English to Arab Students. Okaz Printing Press. Jeddah. KSA.
- Anderson, Scarvia B. : (1975). Encyclopedia of Educational Objectives. Jossey-Bass Publisher San Francisco. California, USA.
- Bande T.M, NKECHI, M. (2008) [online]
- Nordic Journal of African Studies* 17(3): 198–210
- Bloom, Benjamin S (et.al) : (1956). Taxonomy of Educational Objectives. Book 1. Cognitive Domain. Longman Group Limited, New York.
- Bronwyn, Carole: (2003). [Online] at: ([www.cal.org.ericcd/Digest/subject/hhttp#assessment](http://www.cal.org.ericcd/Digest/subject/hhttp#assessment)).
- Lado, Robert : (1975). Language Testing. Wing Tasi Cheung Printing Company LTD. Hong Kong.
- Ebel, Robert: (1972). Essential Educational Measurements. Prentice Hall, Inc. Englewood, Cliff. New York, USA.
- Eric Digest. (2003) (<http://www.ered.com>.)
- Garallah Al- Garallah. (2010). [Online]. [edu.com/inf2/articles.php?action=show&id=210](http://edu.com/inf2/articles.php?action=show&id=210) -
- Hughes, Arthur: (1995). Testing for Language Teachers. Seventh Printing. University of Cambridge. Bell & Bain, Ltd. Glasgow. UK
- Lindquist, E.F.: (1961). Educational Measurement. American Council on Education. Washington D.C. George Banta Publishing Company. Menasha, Wisconsin. USA.
- McNamara, Tim : (2001). Language Testing. Oxford University Press. UK.
- Menke, K in Orlich, Donland.C, (et.al): (1998). Teaching Strategies. 5<sup>th</sup> edition. Houghton Mifflin Company, Boston New York. U.S.A.
- National Centre for Higher Education: [www.qiyas.org](http://www.qiyas.org)
- Nunan, D. and C. Lamb: (1996). [Online], at <http://www.exchanges.state.govern./forum/vols/vol37/no1/index.html>
- Sax, Gilbert: (1980): "Principles of Educational and Psychological Measurement and Evaluation". 2<sup>nd</sup> edition, Wadsworth Publishing Company. Belmont, California. USA.
- Siddiek, Ahmed Guma'a: (2001): " Analysis & Evaluation of Sudan School Certificate English Examination". (Partial M.ED), Juba University. Khartoum, Sudan.