

## **AN INVESTIGATION OF JORDANIAN EFL TEACHERS' ASSESSMENT PRACTICES**

*Nayel Darweesh Al-Shara'h\**

### **ABSTRACT**

The study aimed at investigating Jordanian EFL teachers' assessment practices relating to the test construction through self-reported frequencies of using the procedures of preparing, correcting, analyzing, interpreting an achievement test, and discussing its results with students. To achieve this, a 31-item questionnaire was used. The questionnaire was administered to 118 basic stage EFL teachers after establishing its validity and reliability.

The results showed that EFL teachers claimed to always or usually practice appropriate procedures of preparing the test, discussing the results with students and evaluating and assessing short-answer tests. However, they were found to sometimes practice appropriate procedures for analyzing test results and evaluating and assessing open-ended questions.

In light of the findings, it is recommended that educational institutions should pay more attention to educating teachers to analyze and interpret test results, and evaluate and score open-answer questions.

### **1. BACKGROUND**

Classroom assessment plays an essential role in the teaching/learning process. Through assessment, the teacher can judge the extent to which he/she has achieved his/her planned instructional objectives. And based upon this judgment, he/she decides whether to continue the instructional process, or simply change teaching or instruction in order to address what has not been achieved. Thus, the evaluation process goes on from one lesson to another, making sure objectives are realized.

Classroom assessment uses a range of tools that contribute to making decisions about student achievement. These tools are in two main categories: traditional testing procedures such as multiple-choice, matching, true-false, short-answer and essay tests; and alternative assessments such as observation, conferences, portfolios, peer and group assessment techniques (Aschbacher, 1994; Davies, 1999; Rose, 1996; Genesee & Upshur, 2004). Despite the

---

\* Faculty of Educational Sciences, University of Jordan. E-mail: sharah@ju.edu.jo

existence of both forms of assessment, schools use testing as the major method of evaluation. Jordanian schools, for instance, weight total grades 70% tests and 30% alternative assessments (Ministry of Education 2006 a).

Bose (2003) urges that teachers should assess and test what they teach because the objectives of testing are the same as those of teaching. In other words, the test should test the language skills and the language elements they have been taught to students. He adds that teaching and testing are like the two sides of a coin; one without the other is not useful. Hughes (1989) refers to the effect of testing on teaching and learning as backwash. Backwash can be harmful or beneficial. He adds that everything should be done to maintain the good quality of testing and to improve its practice. Richards (1990) says that if exams are well designed and properly used, they can effectively enhance the educational process. This is because educators see tests as motivators that stimulate individuals to do their best. Further, tests are a means of obtaining systematic evidence on which we can base instructional decisions.

The strong emphasis on tests might be attributed to teachers believing that tests influence students' learning through fostering student motivation and encouraging them to review what they have learned. Meanwhile, tests provide teachers with good feedback of the positive and negative aspects of their instruction (Morgan 2008). However, although tests are very common amongst school teachers, they still have problems and difficulty constructing them properly, according to the appropriate procedures of test construction.

Teachers ought to pay much attention to the way they assess their students' achievement. For example, if they want to use a particular type of test, the assessment should be constructed in a way that guarantees its capability of measuring the extent to which objectives are clearly and practically achieved. Further, the test is expected to represent the content of curriculum, reflecting a good table of specifications, a table of two dimensions, one represents the content, the other represents the thinking processes expected from students (Griswold, 1990; Gronlund, 1998; Vos, 2000). Teachers, on the other hand, should follow the appropriate procedures of test construction that ensure validity and reliability of test: identifying the purpose of the test; good planning and preparation such as determining the intended learning outcomes, selecting a representative sample of items and writing the items in a way that does not have any ambiguity; test administration; good correction; and good analysis of test results (Chittenden, 1991; Kellaghan and Madaus, 1991; Brown, 1998).

Popham (1995) and Daniel and Deber (1998) indicated a number of characteristics with respect to evaluation teachers should have in order to be effective in their classes, amongst which are: knowing how to design and construct a good, educational test; and being able to diagnose the weaknesses and strengths of their students and to follow up their students' learning progress. This is also supported by stiggins, Arter, Chappuis and Chappuis

(2006) who agree that the identification of both the strengths and weaknesses of student performance is essential for effective use of assessment results.

It is obvious, then, that preparing a valid, reliable test requires an assessment literate teacher, who is well-trained and well-prepared to consider the procedures of test construction and implement such procedures in his/her tests. Therefore, educational institutions have thought about competences of measurement and evaluation to be included in teacher education programs, both at the pre-service and in-service levels; as for the evaluation process to succeed, and for teachers to develop professionally, we should provide and empower them with the essential skills of evaluation (Joint Committee on Standards for Educational Evaluation, 2002; Daniel and Deber, 1998; Fitt, Rafferty, Presner and Heverly, 1999). Jordan's Ministry of Education (2006) published a list of standards and levels of performance teachers have to demonstrate in order to develop professionally. One of the main domains considered in this document is assessment of students' learning and instruction. For instance, teachers should demonstrate understanding of linkage between assessments, instruction and learning outcomes. They are also required to choose and design varied and appropriate tools and means for assessing student learning and progress, and they should also analyze students' performance and provide them with feedback about their learning and progress (Ministry of Education, 2006 b).

## **2. STATEMENT OF THE PROBLEM**

The problem of the present study came as a result of the researcher's following up University of Jordan student teachers of English who were doing their practicum in local schools in the City of Amman. During classroom discussions with those student teachers about assessment of student learning, the researcher felt that they had problems with the way they assess their children. Most of their assessment centers around testing, achievement tests in particular. It is true that achievement tests are an essential tool of assessing students' progress. However, we would find that both established teachers and student teachers disagree about the right procedures that should be followed in preparing such tests. The same problem appeared to apply to cooperative teachers who shared university faculty members in educating student teachers. Both cooperative teachers and student teachers appeared to have problems with constructing, administering, correcting, and analyzing achievement tests. Hence, the present study aims to explore the common assessment practices relating to the test construction and the degree to which EFL teachers practice appropriate procedures of the achievement test construction.

To achieve the above aim, the following questions were addressed:

1. To what degree do Jordanian EFL teachers practice appropriate procedures of the achievement test preparation?

2. To what degree do Jordanian EFL teachers practice appropriate procedures of evaluating and assessing short-answer and open-answer questions of the achievement test?
3. To what degree do Jordanian EFL teachers practice appropriate procedures of discussing the achievement test results with students?
4. To what degree do Jordanian EFL teachers practice appropriate procedures of analyzing the achievement test results?

### **3. METHOD AND PROCEDURES**

#### *3.1. Participants*

One hundred and eighteen (118) teachers, teaching English as a foreign language to basic stage students in the University District, Amman, participated in the present study. While 65 of them were male, the rest 53 were female teachers. Regarding experience, the participants were distributed in four groups: 37 had less than 5 years of experience, 27 had experience between 5-9 years, 29 had experience between 10-14, and 25 had more than 14 years.

#### *3.2. Research instrument and data collection*

One data collection instrument was used to provide data for the present study: a *questionnaire* in which participants in the study had to respond to a five-point frequency scale: Always (5), Usually (4), Sometimes (3), Rarely (2), and Never (1). This instrument was adopted with some modification from Al-Younes (2006). The questionnaire comprised (31) items, distributed in five dimensions as follows: preparing for the test (6) items; evaluating and assessing the test whose answers are closed (6) items; evaluating and assessing the test whose answers are open (3) items; discussing the test results with students (7) items; and analyzing test results (9) items.

The questionnaires were sent to the participating teachers via student teachers who were doing their Practical Education course in the nearby schools, or via graduate students who were teachers of English and doing their M.A. degree in TEFL in the Faculty of Educational Sciences at the University of Jordan. The questionnaires were filled out and sent back to the researcher.

#### *3.3. Validity and reliability of the research instrument*

Although the instrument was adopted from Al-Younes (2006), it was also given to a panel of five judges, two of them specializing in curriculum and

instruction and three in measurement and evaluation. Most of their comments focused on the wording of the items in the different dimensions of the questionnaire. All their comments were taken into consideration in the final version of the questionnaire.

The internal consistency of the five dimensions of the questionnaire was computed using Cronbach Alpha coefficient. Table (1) shows the dimensions, number of items in each dimension and Cronbakh Alpha coefficient.

**Table (1)**  
**Cronbach Alpha coefficients for each dimension of the questionnaire**

| No. | Dimension   | No. of items | Reliability |
|-----|---|--------------|-------------|
| 1.  | Preparing for the test                                    | 6            | 0.72        |
| 2.  | Evaluating and assessing the test whose answers are short | 6            | 0.87        |
| 3.  | Evaluating and assessing the test whose answers are open  | 3            | 0.83        |
| 4.  | Discussing the test results with students                 | 7            | 0.84        |
| 5.  | Analyzing test results                                    | 9            | 0.83        |

#### 4. FINDINGS AND DISCUSSION

As noted above, a five-point Likert scale was used in the questionnaire of the present study: Always (5), Usually (4), Sometimes (3), Rarely (2), and Never (1). In order to make it easier to interpret the findings of the study, the following criteria (out of 5) were adopted to judge the degree to which the procedures of preparing, correcting, discussing and analyzing the test were appropriately practiced by participant teachers: 1.0-1.49= never practice, 1.50-2.49= rarely practice, 2.50-3.49= sometimes practice, 3.50-4.25= usually practice, 4.26-5.0= always practice.

To answer Question One: *To what degree do Jordanian EFL teachers practice appropriate procedures of achievement test construction?* Means and standard deviations were computed for each item in the four dimensions of the questionnaire. In what follows is a presentation of the findings:

##### 4.1. Findings relating to test preparation

This dimension consisted of 6 items. Table 2 shows the means and standard deviations of the degree to which participant teachers practice the procedures of achievement test preparation.

**Table (2)**  
**Means and standard deviations of the participant teachers' responses to the degree of practicing the procedures of preparing for the test**

| No. | When preparing for the test, I consider:  | Mean | Std. Deviation |
|-----|---|------|----------------|
| 1.  | determining the purpose of the test   | 4.55 | .74            |
| 2.  | test representation of the cognitive domain levels  | 4.27 | .77            |
| 3.  | determining the type of test or questions (Yes/No, Matching, Multiple Choice, Short-answer, Essay, etc.) to be used in the test | 4.30 | .84            |
| 4.  | determining the number of items or questions to be used in the test   | 4.28 | .92            |
| 5.  | representation of test questions to the objectives and content  | 4.46 | .77            |
| 6.  | the importance of the topic by asking an appropriate number of questions/items  | 4.35 | .75            |
|     | <b>Total</b>  | 4.37 | .80            |

Further, the standard deviations, as noted in the table, indicate a high degree of agreement and consistency in the participant teachers' responses to all items of this dimension of test achievement construction.

Examining the above findings, we notice a high degree of awareness amongst teachers regarding the importance of planning and preparing the achievement test. And according to the suggested criteria identified in our scale above, we find our teachers always consider all the procedures of test preparation. Generally speaking, the findings achieved in this domain are pleasing, indicating that teachers' practices are adequate and in harmony with the new standards required by the Ministry of Education in Jordan and the educational institutions worldwide.

#### *4.2. Findings relating to evaluating and assessing test questions*

This dimension comprised two parts: one asking teachers about what procedures they use when correcting a short-answer question, the other is concerned with procedures followed when correcting an essay question.

##### *4.2.1. Findings relating to evaluating and assessing short-answer questions*

This part of the questionnaire consisted of 6 items. Table 3 shows the means and standard deviations of the degree to which participant teachers practice the procedures of correcting an achievement test with short-answer questions.

**Table (3)**  
**Means and standard deviations of the participant teachers' responses to the degree of practicing the procedures of correcting the short-answer test**

| No. | When correcting the short-answer test, I consider:  | Mean | Std. Deviation |
|-----|---|------|----------------|
| 1.  | When the test includes more than one question, I correct one question at a time before I move to the next.                    | 3.29 | 1.38           |
| 2.  | I prepare a modal answer for each question.   | 4.24 | 1.00           |
| 3.  | I distribute the marks according to all points mentioned in the typical answer.   | 4.49 | .75            |
| 4.  | I am committed to the points mentioned in the modal answer, considering any possible modifications.                           | 4.11 | .94            |
| 5.  | I already assign a certain weight of the grade that matches the importance of the content and the objective that measures it. | 4.39 | .80            |
| 6.  | I correct all the questions in each student's paper, then I move to the next.   | 3.87 | 1.22           |
|     | <b>Total</b>  | 4.07 | 1.02           |

Further, the standard deviations, as noted in the table, indicate a high degree of agreement and consistency in the participant teachers' responses to Items 3, 4 and 5 of this dimension of test achievement construction. The standard deviations of Items 1, 2 and 6, however, indicate a discrepancy and relatively clear variation in the respondents' answers, especially with respect to: *correcting one question at a time, before moving to the next*, and *correcting all the questions in each student's paper, then moving to the next*, Items 1 and 6, respectively.

Examining the means scored in the participants' responses to correction procedures of achievement tests when evaluating short-answer questions, we can see that the means are high with regard to Items 3 and 5, which means that teachers *always* practice the procedures of *distributing the marks according to all points mentioned in the typical answer* (4.49), and *commitment to the points mentioned in the modal answer, considering any possible modifications* (4.39). This of course shows a high degree of awareness amongst teachers regarding such procedures. However, according to the suggested criteria of our scale above, we find our teachers *usually* consider the other procedures of this dimension. Generally speaking, the findings achieved in this domain are satisfactory and reveal our teachers' awareness of the importance of applying appropriate procedures when

evaluating and assessing their students' short-answer tests, which again responds to the standards required by educational institutions.

#### 4.2.2. Findings relating to evaluating and assessing open-answer questions

This part of the questionnaire consisted of 3 items. Table 4 shows the means and standard deviations of the degree to which participant teachers practice the procedures of evaluating an achievement test with open-answer questions.

**Table (4)**  
**Means and standard deviations of the participant teachers' responses to the degree of practicing the procedures of evaluating the open-answer test**

| No. | <b>When correcting the open-answer test, and it is difficult to analyze it to some major points, I do the following:</b>     | Mean | Std. Deviation |
|-----|--|------|----------------|
| 1.  | I read all students' answers before starting to give a grade.  | 3.75 | 1.07           |
| 2.  | I classify the students' papers in grades or categories (A, B,C, D, etc.)  | 2.77 | 1.13           |
| 3.  | I reread the students' answers in each group or category and move some of them to other categories according to performance. | 3.12 | 1.24           |
|     | <b>Total</b>   | 3.21 | 1.15           |

As is seen in Table 4, the mean scored by respondents was somewhat high with regard to Item 1 of this dimension: *reading all students' answers before starting to give a grade* (3.75). The means of the other two items (*classifying students' papers in grades or categories (A, B, C, D, etc.,* and *rereading the students' answers in each group or category and move some of them to other categories according to performance* ) were somewhat low (2.77 and 3.12, respectively). What is more, the standard deviation of Item 1 showed a sort of harmony in the participant teachers' answers. This, however, is not reflected in Items 2 and 3, whose standard deviations indicated a discrepancy in the teachers' responses.

Looking at the findings of this part, one can note that the teachers' performance on the procedures of evaluating an achievement test with open-answer questions is not very satisfactory, which may lead to inaccuracy of evaluation and lack of objectivity. This finding is consistent with that of Al-Younes (2006), which also reported unsatisfactory performance of respondents on the procedures of evaluating essay tests.

4.3. Findings relating to practicing the procedures of discussing the test results with students

This part of the questionnaire consisted of 7 items. Table 5 shows the means and standard deviations of the degree to which participant teachers practice the procedures of discussing the test results with students.

**Table (5)**  
**Means and standard deviations of the participant teachers' responses to the degree of practicing the procedures of discussing the test results**

| No. | When returning test results to the students, I do the following:   | Mean | Std. Deviation |
|-----|--|------|----------------|
| 1.  | I answer all the test questions to the students, either on the chalkboard or on a separate sheet.          | 4.21 | .93            |
| 2.  | I explain to students how I distributed marks to answers.  | 4.30 | .87            |
| 3.  | I give a chance for students to discuss their answers and review their results.                            | 4.69 | .61            |
| 4.  | I accept students' dialogue and defense for their answers.   | 4.64 | .62            |
| 5.  | I accept students' enquiries about the way used in rating and correcting their answers.                    | 4.31 | .88            |
| 6.  | I show the distribution of students' results graphically, or show them in a descending or ascending order. | 2.29 | 1.31           |
| 7.  | When necessary, I write some comments on students' answers.  | 3.58 | 1.14           |
|     | <b>Total</b>   | 4.28 | .98            |

As shown in Table 5 above, the means scored regarding the participants' discussing the test results with their students range between 2.29 and 4.69. The highest means were scored by Items 3, 4, 5 and 2: 4.69, 4.64, 4.31 and 4.30, respectively. Thus, the participant teachers were found to *always give a chance for students to discuss their answers and review their results, accept students' dialogue and defense for their answers, accept students' enquiries about the way used in rating and correcting their answers, and explain to students how they distribute marks to answers*. Relatively high means were also scored as regards Items 1 and 7: *answering all the test questions to the students, either on the chalkboard or on a separate sheet* (4.21), and *writing some comments on students' answers, when necessary* (3.58). On the other hand, we notice a low mean scored with respect to Items 6: *showing the*

*distribution of students' results graphically, or show them in a descending or ascending order (2.29).*

A quick look at the standard deviations of Items 1, 2, 3, 4 and 5 show a high degree of consistency in the responses of the sample of the study, unlike Items 6 and 7, which indicate discrepancy and lack of harmony in the participants' responses.

The findings above reveal that EFL teachers always practice most of the procedures included in *discussing achievement test results with their students*. However, scrutinizing the findings relating to Item 6, one can note teachers' low performance on *showing the distribution of students' results graphically, or showing them in a descending or ascending order, especially when we know how important it is for students to know their level of performance in contrast with their classmates*. This again reinforces the teachers' transparency and objectivity with his/her students (Brown, 1998; Al-Younes, 2006).

#### 4.4. Findings relating to practicing the procedures of analyzing test results

This section of the questionnaire consisted of 9 items. Table 6 shows the means and standard deviations of the degree to which participant teachers practice the procedures of analyzing test results with students.

As is apparent in Table 6, except for Items 3 and 9, we can see the low performance of participant teachers on most of the procedures that comprise this domain of achievement test construction, i.e. test results analysis. While teachers appear to *usually* practice the procedures of: *computing the percentages of pass and fail, and keeping the good items of the test to use them in the future*, they do not perform well the other procedures of test results analysis such as *computing the mean, median, and mode that indicate the centeredness of students' marks, and computing other necessary statistical devices such as standard deviations, difficulty coefficients, and discrimination coefficients*, which give the teachers significant information about the test and student performance on it. Further, standard deviations show a high level of disagreement and inconsistency in participants' responses with respect to all items of this dimension, except for Items 3 and 9 which reflect a sort of consistency in responses.

Examining the findings of this dimension against the suggested criteria of our scale above indicates EFL teachers' lack of awareness of these important procedures, which also indicates that the concept or *culture* of test results analysis is almost lacking in our educational settings. This could be ascribed to the teachers' limited knowledge of statistics. These findings are staggering, especially when we know how significant analyzing the test results is. Test results analysis makes the teacher confident about his test items, and this in turn makes him/her benefit from such test items in the future, and this consequently contributes to his/her ability to improve his/her tests, and

making them valid, practical and reliable (Brown,1998; Gronlund, 1998; Daniel and Deber, 1998; McMillan, 1999; Al-Younes, 2006).

**Table (6)**  
**Means and standard deviations of the participant teachers' responses to the degree of practicing the procedures of analyzing the test results**

| No. | When analyzing test results, I do the following:  | Mean | Std. Deviation |
|-----|---|------|----------------|
| 1.  | I arrange the test papers in an ascending or descending order.                                      | 3.21 | 1.43           |
| 2.  | I compute the mean, median, and mode that indicate the centeredness of students' marks.             | 3.26 | 1.42           |
| 3.  | I compute the percentages of pass and fail.   | 4.08 | 1.14           |
| 4.  | I compute the standard deviation that indicates the extent to which marks are scattered from means. | 2.86 | 1.41           |
| 5.  | I compute the difficulty coefficients of the items or/and questions.                                | 3.11 | 1.30           |
| 6.  | I compute the discrimination coefficient for each item  | 2.55 | 1.31           |
| 7.  | I compute the efficiency of distracters when I use a multiple-choice test.                          | 2.73 | 1.33           |
| 8.  | I compute the percentage that comes below or above each student's score.                            | 2.78 | 1.33           |
| 9.  | I keep the good items of the test to use them in the future.  | 4.01 | 1.07           |
|     | <b>Total</b>  | 3.18 | 1.30           |

In conclusion, examining the total means of the major dimensions of the study (Tables 2-6), we can notice that these means range between 3.18 and 4.28. According to the adopted criteria of the scale used in the study, teachers appear to *always* or *usually* practice the appropriate procedures of test construction with respect to the three dimensions: *discussing the test results with students* (4.28), *evaluating and assessing short-answers tests* (4.07), and *preparing for the test* (3.87). Such findings are consistent with what has been indicated in Daniel and Deber (1998), Gronlund (1998), and Griswold, (1999). As regards the other two dimensions, teachers seem to just *sometimes* consider in their practices: *evaluating and assessing open-answers tests* (3.21) and *analyzing test results* (3.18). These findings come opposite to what Griswold (1999), MacMillan, (1999), Sanders (2001), and Al-Younes (2006) call for. Such findings do not reflect the advantages hoped for from the

process of analyzing test results, for instance, which represents good feedback for teachers. Furthermore, objective evaluation and assessing of students' answers is a very necessary practice teachers are to be aware of (Griswold, 1999; and Sanders, 2001).

## 5. RECOMMENDATIONS

In light of the above findings, it is recommended that educational institutions pay more attention for qualifying teachers with respect to analyzing and interpreting test results, and evaluating and assessing tests with open-answer questions. The Ministry of Education is invited to hold more teacher education courses on assessment and evaluation, with more focus on test results analysis. The Ministry of Education is also invited to include the procedures of achievement test construction in the booklet of *National teacher professional standards*. Further research on other assessment practices that teachers are required to use to assess student learning is also recommended.

## REFERENCES

Aschbacher, P.R. (1994). Helping educators to develop and use alternative assessments : Barriers and facilitators. *Educational Policy*, 8 (2), 202-23.

Bose (2003). A letter to teachers of English: 32 Test what you teach (1), Yemen Times. Retrieved April 12 2004 from <http://www.yemintimes.com/print.article.shtml?i=education&a=2>.

Brown, W.R. (1998). Test mapping: Planning the classroom test. *Clearing House*, 71 (3), 153-54.

Chittenden, E. (1991). Authentic assessment, evaluation and documentation of students performance. In V. Perrone (Ed), *Expanding student assessment* (pp. 22-31). Alexandria: Association for Supervision and Curriculum Development.

Daniel, L.G., & Debra, A. (1998). Knowledge and use of testing and measurement literacy of elementary and secondary teachers. *Journal of Educational Research*, 91 (6), 331-44.

Davies, M.A. (1999). Alternative assessment: New directions in teaching and learning. *Contemporary Education*, 71 (1), 39-45.

Fitt, D., Rafferty, K., Presner, M.T. & Heverly, M.A. (1999). Improving the quality of teachers' classroom tests. *Education*, 119 (4), 643-8.

Genesee, F. & Upshur, J.A. (2004). *Classroom-based evaluation in second language education*. Cambridge University Press.

Griswold, P.A. (1990). Assessment relevance and reliability to improve the quality of teacher made tests. *NASSP Bulletin*, 74( 523),18-24.

Gronlund, N.E. (1998). *Assessment of students achievement*, (6th Ed.). MA: Allyn and Bacon.

Hughes, A. (1989). *Testing for Language Teachers*. Cambridge: Cambridge University Press.

Joint Committee on Standards for Educational Evaluation. (2002). *The student evaluation standards: How to improve evaluations of students*. Thousand Oaks, CA: Corwin Press.

Kellaghan, T., & Madaus G.F. (1991). National Testing: Lessons for America from Europe. *Educational Leadership*, 49 (3), 87-93.

McMillan, J.H. (1999). *Establishing high quality classroom assessments*. Metropolitan Educational Research Consortium. (ERIC Document Reproduction Service No. ED429146).

Mertler, C.A. (1999). Teachers' Misconceptions of classroom test validity and reliability. Paper presented at the Annual Meeting of the Mid-Western Educational Research Association (Chicago, IL., October, 13-16). (ERIC Document Reproduction Service No. ED437349).

Ministry of Education (1990). *English language curriculum and its general guidelines for the basic education stage*. Jordan.

Ministry of Education. (2006 a). *General guidelines for curriculum and evaluation*. Jordan.

Ministry of Education (2006 b). *National teacher professional standards*. National teacher professional development standards conference. Amman May 16-18 2006.

Morgan, A. 2008. *Feedback: Assessment for rather than of learning*. Retrieved September 3, 2008 from [http://www.bangor.ac.uk/adu/the\\_scheme/documents/FEEDBACKJanuary06\\_000.ppt](http://www.bangor.ac.uk/adu/the_scheme/documents/FEEDBACKJanuary06_000.ppt)

Oosterhof, A. (1996). *Classroom applications of educational measurement*, (2nd. Ed.). NY: Merrill.

Popham, W. J. (1995). *Classroom assessment: What teachers need to know*. Needham Heights: Allyn and Bacon.

Richards, J.C. (1990). *The Language Teaching Matrix*. New York: Cambridge University Press.

Rose, E.W. (1996). The role of portfolio evaluation in social studies teacher education. *Social Education*, (60), 162-166.

Sanders, L.R. (2001). Improving assessment in university classroom. *College Teaching*, 49 (2), 62-4.

Stiggins, R., J. Arter, J. Chappuis, & Chappuis. 2006. *Classroom assessment for student learning: Doing it right – Using it well*. Portland, OR: ETS Assessment Training Institute.

Vos, H. (2000). How to assess for improvement of learning. *European Journal of Engineering Education*, 25 (3), 227-233.

Wiersma, W. & Jurs, S.G. ( 1990). *Educational measurement and testing*, (2nd. Ed.). Massachusetts: Allyn and Bacon.

Al-Younes, Y.A. (2006). The extent of the University of Jordan faculty members' practices of test preparation, correction, analysis, interpretation, and

discussion of results with students. Assyout University *Journal of the Faculty of Education*, 22 (1), 153-193.