

The Effectiveness of Evaluation Style in Improving Lingual Knowledge for Ajloun College Students

Mohammad Fawzy Bani-Yaseen, Mohammad Akram Al-Zu'bi, Haytham Abu-Zaid
Al- Balqa Applied University, Ajloun University College.

Abstract

The study aimed at investigating to what extent does evaluation style help in improving the lingual knowledge for females' students in Ajloun College. The sample of the study consists of (45) students from Ajloun College and it was selected purposefully; it represents three experimental groups; the first group consists of (15) students was evaluated by the teacher without discussing the test after evaluation; the second group consists of (15) students was evaluated by classmates; and the third group consists of (15) students was assessed by themselves. The treatments were distributed on the groups randomly. The results revealed that there are significant statistical differences at ($\alpha \leq 0.05$) among the performance means of the post test of the groups. Thus, the post comparisons indicated that there were significant differences at ($\alpha \leq 0.05$) among the performance means due to the group that was assessed by classmates. Also, there were significant differences at ($\alpha \leq 0.05$) between the group that assessed by the teacher and the group that assessed by themselves in favor of the group that was assessed by themselves.

Key Words: evaluation style, lingual Knowledge.

Preface

Obeidat and Abu-Alsamen (2005) describes the process of evaluation as a mental skill associated with the ability of making judgments, decision-making, the accuracy of their sources, and detection of inaccuracies and errors contained in the opinions and ideas. Also, The Analysis evaluation considered as successful method to develop language skills (Ministry of Education, 2005a). The evaluation process of the effective classroom should be holistic, continued, economical, collaborative, democratic, scientific, flexible, fair, realistic, and meaningful.

Abu Hashim and Alsaid (2010) presented differences among related terms, as follows: the testing is a final process measures one of the student aspects, or a regulator to measure a trait of a sample of behavior, while evaluation is extended to include different aspects of the student's in order to give a picture of the growth of these respects. Also, the testing carried out by only the teacher who designs it then determines the time and place to apply and correct it but evaluation is a comprehensive cooperative process involving everyone related to the instructional process. Testing is a standardized process measures the adequacy of the individual in one of the areas but evaluation is a therapeutic process that diagnoses the current situation, it does not stop at this stage, but extend to place the appropriate treatment. Measurement is prior to evaluation and the basis for it and refers to the amount of what existing in the thing according to the standards listed with a numeric value agreed them but evaluation is limited to sentencing of the things value, in other words to assess the relationship between the level of achievement and goals (assessing the value of something based on a certain standard) and it means to make value judgment on the result of measurement according to the standard theme.

The evaluation shall be deemed a broader of these terms and the most comprehensive. It is defined as a systematic process based on scientific principles, aimed at the sentencing accurately and objectively on the inputs, processes and outputs of any educational system, and then determine the strengths and shortcomings in both of them, in order to take appropriate decisions to fix what might be revealed from weaknesses and deficiencies. Each evaluation process includes the processes of measurement and assessment, and it is not necessarily depending on the process of measurement. Measurement and evaluation processes alone are not singing about the evaluation process. The evaluation process is an ongoing process which is present in every situation and take the various parts of the question and the answer, solving exercises. It is also a comprehensive process of learning and teaching process so that it displays every position in education. It is better to invoke the teacher in the process of teaching (Abu Amsha, 2011).

Background and Theoretical Framework

Evaluation is divided into two strategies: verified written tested evaluation which is the most widely used strategies. The other division is unwritten tested evaluation which is an alternative evaluation that the teacher can determine if he understands the educational learning performances of his students. According to Fadlallah (2005) that the alternative evaluation derives directly from the standards and associated with Tail reference evaluation. In the opinion of Sa'aadah and Abraham (2004) that the alternative evaluation is a set of information, data and judgments that help in reviewing and developing instructional programs based on firm foundations to evaluate the quality of instruction. According to Al-Rubaie (2006a), the good evaluation is to take advantage of the

collected information and use strategies of non-traditional written tested evaluation. The alternative evaluation strategies considered as a way for real learning, and give an opportunity for critical thinking and knowledge building (Muirhead, 2002).

Because of the multiple forms of alternative evaluation, the teacher knows how to achieve the expected goals. Taha (2006) confirms the need to exercise in the alternative evaluation. As Anderson (2002) has stressed the need for practicing alternative evaluation to check the extent to which the objectives of the language in the basic language skills (reading, writing, listening and speaking) achieved. Also, (Yates, 2000) believes that students in this type offering their cognitive skillful performances and it gives them adequate opportunity to perform the produced performances and employed them.

On the other hand the measurement means to express the degree of student ownership of a particular attribute in numbers. The tools of measurement are the tests and observation, interview and records.

The linguistic evaluation areas can be displayed as follows:

- The field of the linguistic adequacy (grammar, order system, morphological and semantic system, the rules and laws that govern the work of language, and related concepts).
- The field of linguistic performance (listening, reading, speaking and writing), and the mental skills associated with these skills.

A balances between the criterion evaluation and referenced could be found as follows:

Criterion evaluation

- Test scores are more holistic and comprehensive.
- Scores considered as a unit of measurement and the balance.
- Explaining scores by Return them to limited standards.
- Writing test items by taking into account the type of performance.
- Re-test needs to the accuracy and skill that may not be available sometimes in the teachers.
- The standard is a group.

Referenced evaluation

- Test scores directed towards a particular goal.
- You get a large number of degrees because of the multiplicity of objectives.
- The teacher defines the points and degrees of the pieces, and level of the limited efficiency.
- The elements of the test are identified by the achieved objectives learning situations.
- It is easy to interpret the test results by teachers after setting goals and building test
- Setting goals to be achieved.
- Focused on the objectives and outputs (Qatami, and Abu-Jaber and Qatami, 2002)

The Assessment Strategies could be presented as follow:

1. Performance- based assessment strategy: includes presentation, demonstration, performance, speech, exhibition, simulation, debat.
2. Pencil and Paper.
3. Observation.
4. Communication(Interview, Conference).
5. Reflection which contains self-evaluation, students' diaries, and students files (Ministry of Education, 2005b).

Odeh and malkawi (2001) have mentioned a set of conditions that must be available in the tools of evaluation which are validity, reliability, test-wash back, varity.

The test should be according to the following steps: identifying the purpose of the test, analyzing the content, building a table of specifications, designing test, applying testing, analyzing test paper, and correcting of the test. . Also, Good test takes into account the time factor; time must be enough for the average student, the test fails if time is not enough, it is known that the error in estimating the time is two types: increasing error and decreasing error, and the first error is less harmful than the second error (Al-Kholi, , 2000). The language tests aimed to assess the language proficiency of the language learner that represents his linguistic performance in meaningful contexts and communicative alive situations (Abu Hija, 2005).

The attention and focus on the evaluation of performance became in response to direct sharp criticism made against the essay and objective tests as traditional standard (Sabri and Rafii, 2001; Al-Faouri, 2006), and the new trends in the field of educational evaluation confirmed alternative evaluation; It is based on the assumption that knowledge is being created and built by the learner; as that knowledge is different from one context to another, and the alternative evaluation based on the authentic learning tasks that teach students works facing the adult in their field (Wiggins, 1998). According to (Antoinette, 1996), the alternative evaluation requires that students should apply what they learn and perform a variety of tasks and actions. Also, (Abu Alam, 2001) confirms the need to realistic confirms, as well as the need to apply the knowledge and skills wisely and effectively to solve

unorganized problems, to give opportunities for training, and to reflect a shift from the view of the consignment view of learning to constructivist learning (Al-Sarraf, 2002). Mahernz and Lehman (2003) have mentioned many reasons that make performance evaluation the best form of educational evaluation forms; including: the use of realistic representations, diversity in the forms of the evaluation, and objectivity.

A portfolio is an effective tool of alternative evaluation forms, the importance of portfolios has confirmed by: (Allam, 2003; Kwafhah, 2003; Al-Sarraf, 2002; Abdel-Hamid, 2002; Sabri and Rafii, 2001; Al-Dosari, 2000; Allam, 2000; Melhem, 2000; Al-Rousan, 1999; Altreri, 1997; Antoinette 1996). Asp, (2000) and Kelly (2000) who urge to use the technology of computers; as they make the evaluations more elaborate and appropriate to the needs of each individual learner and teacher alike.

Evaluation multiple measures is a type of evaluation that does not rely on one indicator or one method of measurement in judging on the learner, but relies on several techniques, and more than one indicator to judge the level of any element of the inputs, processes and outputs of the educational system (Asp, 2000). Allowably (2003) suggests three main phases experienced by the multiple measurements: the stage of evaluation and diagnosis, the stage of classification and description, and the stage of constructing a model for the required needs. Abd-Alkhaleq (1996) has mentioned a set of language tests which able to achieve a balance between traditional tests and qualitative tests; like the audio- understand tests, Image and words test, testing of the picture and sentences, images and sentence test, different word test, testing of the identical words, testing of different sentence, testing of the identical sentences, testing of heard word, the test of sentence type, testing of the sentence and meanings

Related studies

The studies were listed according to the chronological order (from newest to the oldest), and the display was ended by commenting to show the extent of benefit, and highlight the most important features of the current study as follows:

Alhayajnah (2007) conducted a study aimed at recognizing to what extend are the teachers of Arabic use the alternative evaluation strategies and the effect of training course on their behaviors of those strategies. The population of the study is teachers of Arabic who teach in the secondary stage in First Irbid Directorate. The population consisted (155) but the sample was (18) and selected randomly (18 female teachers and 12 male). The participants divided into two groups: experimental one consists of (15) and controlled of (15). The findings of the study revealed that there are statistically significant differences in the use of Arabic teachers the alternative evaluation strategies in favor of the experimental group and there are statistically significant differences due to the experience and no statistically significant differences due to the interaction between the program and the experience.

Allen and Flippo (2002) examined attitudes and concerns related to the use of alternative assessments in literacy education courses. A Likert-style twenty-item questionnaire was developed and administered first as a pilot. Three subscales were identified: Self-Evaluation, Peer Evaluation, and Instructor Modeling. The questionnaire was then administered to two sections (one literacy assessment course as pre- and post-course assessments). The results were significant for Self-Evaluation and Peer Evaluation. Results for Instructor Modeling did not reach significance but were generally positive in support of modeling alternative assessment in classes.

Nakamura (2002) conducted a study aimed at examining teacher and peer evaluation of English language oral presentation skills in Japanese classrooms. Twelve graduate university students' oral class presentations were assessed by a teacher and four peers. After presentations were rated, teachers calculated the evaluations, and classmates shared their opinions on the top of three presenters. The teacher gave summary comments on all presenters. The Many-Facet Rasch Measurement Model was used to investigate three areas (student ability, item difficulty, and rater severity). The study indicates that peer assessment can successfully motivate students to improve their presentations, students can be reliable peer raters, and the Rasch model can yield considerable significant data on several factors involved in presentation assessment.

EL-Koumy (2001) investigated the effects of student self-assessment on English-as-a-Foreign-Language (EFL) students' knowledge achievement and academic thinking. Ninety-four college seniors from the Department of English at the School of Education at Suez University in Egypt participated in the study. They were randomly divided into two groups, an experimental group and a control group. In the experimental group, throughout the semester, students were asked to independently assess their own knowledge and thinking before and after each lecture. In the control group, students were taught the same English language teaching methodology course with the same method without self-assessing of their knowledge or thinking. Prior to and at the end of the semester, all subjects were tested on both knowledge and academic thinking. The data were analyzed using the t-test. It was found that the mean score gain of the experimental group was slightly higher than that of the control group on both knowledge achievement and academic thinking. The difference was not significant at the 0.05 level.

Therefore, there is not sufficient evidence to conclude that self-assessment can improve students' knowledge or academic thinking.

Adel and Barham (2012) investigated a study aimed at investigating the degree of Mathematics and Arabic teachers' using of the alternative assessment strategies and its tools in Jordan. To achieve the objectives of the study; a questionnaire was built and it was distributed over 86 teachers, and semi- structured interviews were conducted with 20 teachers from the two specializations. Results of the study revealed that the degree of teachers who use the pencil and paper strategy was high, while it was intermediate for the use of performance –based assessment strategy, the observation strategy, and the communication strategy, and it was low for the reflection assessment strategy and using the alternative assessment tools. Results also revealed that there were no statistically significant differences in the degree of using the alternative assessment strategies by teachers related to the effect of teachers' specialization. Whereas, there was statistically significant differences related to the effect of number of years of experience and to the effect of the training courses.

Most of previous studies are foreign so it emphasizes the importance of this study. Also, the Arabic related studies focus on assessment and school evaluation but the present study deals with evaluation and university level. Although some of related studies treated some fields of evaluation at university level, the study investigated the effectiveness of evaluation style in improving lingual Knowledge for Ajloun College students.

Problem of the study and its importance

Researchers suggest that the trends of faculty members in Jordanian universities towards analytical evaluation (or Alternate evaluation) are positive, the practices do not reflect the level of these trends, and meet them (Al-Saikhan, 2009; Hayajnah, 2007; A-Shboul, 2004; Allen and Flippo, 2002; Al-Kaddemat, 2001; and Nasr, 1998). When we look at the ways of students evaluation in Jordanian universities, we find that the space of alternative evaluation application is different; as the most universities expect that a faculty member take the first exam (20%), the second exam (20%), participation is (10%) and a final exam is (50%), and this division reflects the space available for the practicing the different assessment forms. Researchers are trying to stand on the appropriateness of certain types of evaluation in the development of the linguistic knowledge; as relevant Arabic studies are few. Although the concentration of research on school education level, the researcher should focus on the university education level and its evaluation so the researcher carried out this study; since Obeidat and Abu-Alsamen(2005) described the evaluation process as mental skill associated with ability to make judgments, make decisions, notice how the information is correct and accurate sources, and disclosure of inaccuracies, and errors contained in the opinions and ideas, which makes the teacher being away from rapid generalization errors or wrong decisions. As a result, most of experts claimed the ineffectiveness of traditional evaluation; including: (Miller and etl., 2009; Marzano and Haystead, 2008; Nitko and Brookhart, 2007;., Rose, 1999, and Pryor, 1998).

Research purpose and its hypothesis

The purpose of the study is to discover the effectiveness of evaluation style in improving lingual Knowledge for Ajloun College students, so the following null hypothesis is tested:

- There are no statistically significant differences at ($\alpha=0, 01$) in the mean scores of the language test due to evaluation style.
- Operational Definitions
- Evaluation style: the evaluation method conducted on the groups (the control group and experimental) one evaluated by peer and one by self- evaluation.
- Language development: The students' ability to put their theoretical knowledge in the language systems (syntax, semantics, dictation) in practice by scores obtained by students.

Limitations of the study

The findings of the study should be limited by the following factors:

- The subjects of the study are restricted to Ajloun University College female students.
- Assessment is limited to students who are studying language skills course.
- The size of sample is limited to (45) students were selected randomly.
- The test is limited to products of language skills course.

Procedures and Methodology

Study subjects

The population of the study consists of Ajloun College students in the academic year (2012). The sample of the study consists of 45 female students that is purposefully chosen and assigned to three sections: The treatments were distributed on the three groups randomly. Fifteen students were chosen randomly by using (cloud and returns).

To test the equivalence in the groups before carrying out the treatment, the researcher carried out pre-test in same condition and mean scores and standard deviation were computed as shown in Table 1, and ANOVA statistical procedure was computed as shown in Table 2.

Table 1
 Mean scores and standard deviation for pre-test according to evaluation style

Group	No.	Means	standard deviations
Control Group	15	45.00	8.69
Self-evaluation	15	43.80	8.07
Peer-evaluation	15	44.13	9.41
Total	45	44.31	8.56

As can be seen from Table 1, the groups mean scores and standard deviations of pre-test were slightly differ due to the evaluation style (control, self-evaluation, and peer evaluation). To show the significance of statistical differences between the mean scores, ANOVA is used according to Table (2):

Table (2)
 ANOVA for the effectiveness of evaluation style in pre-test

Source	Sum of Squares	Degree of freedom	Mean Squares	F	Significance
Between groups	11.511	2	5.756	.075	.928
Within groups	3210.133	42	76.432		
Total	3221.644	44			

As can be seen from Table 2, there are not significant statistical differences at ($\alpha \leq 0.05$) due to the style of evaluation. So, it means that students' performance is equivalent and the beginning point for each group is closely the same to another group.

Instrument

For the purpose of the study one instrument is developed by the researcher which is a test applied on the participants of the study twice (pre/post test) and the period between them was seven weeks. The following is a brief description of the test:

The test was corrected as follows: 25 scores for every linguistic system (syntax, dictation, grammar, and semantics). It includes short answers (true or false), every statement includes behavioral indicator. Finally, the process of correction done on the test as a whole.

Test validity

The test is content validated by a panel of experts. Comments and criticism on the original draft of the test will be solicited from a number of experts from Jordanian private universities as well as Directorate of Examination, local supervisors and teachers in the ministry of education. The researcher followed the comments of judges by editing, deleting and adding then to produce the final draft.

Reliability of the test

Reliability has been verified by applying on exploratory sample from outside of the study sample belonging to its population, The sample consisted of (23) students, the following scientific steps are adopted to verify the reliability of the test, researchers carried out the repeated reliability according to the following: The researchers applied the test on the exploratory sample consisting of (23) female student, re-testing on the sample itself has done after two weeks, and researchers tried to provide the atmosphere itself in the first application, and papers were corrected, and reliability repetition coefficient of the results calculated for both applications, the correlation coefficient was (89.4) and it is significant at ($\alpha \leq 0.05$) And it is suitable for the purpose of the present study.

Study procedures and implementation

The present study carried out in accordance with the following procedural steps: looking at educational literature based on the styles of evaluation. In the light of this study, the problem of the study has been identified, and the hypotheses were formulated, and determine the population of the study, and behavioral indicators of language skills, and determine the length of time needed to conduct the experiment for seven weeks. indicators were presented to the reviewers, building and validating the test to determine the suitability for the purpose of the study, and testing on a sample reconnaissance to check its reliability, and a pre-test to verify the extent of groups equivalence, and application of post-test on the students after they completed the planned duration of the experiment, and correcting test, and enter data into a computer and analyzed to determine the results and discussion.

Study Design

The study has the independent variable that has three levels (control group evaluation, self- evaluation, and Peer assessment) and the dependent variable (the performance in language test)

Statistical treatment

The researcher used the following measures:

1. Means and standard deviations were computed to compare means of the two groups on the pre and post tests of the reading comprehension.
2. Analysis of covariance ANCOVA was computed to detect any significant differences between the two groups on the reading.
3. Scheffe Post Hoc Test for Evaluation Style Variable in improving lingual Knowledge

Results of the Study

After Applying the post-test and analyze its results, these results have displayed according to the study hypotheses and its variables. To verify hypotheses, means and standard deviations were extracted for the performance of students groups in the post-test and Table 3 shows that:

Table (3)
 Means and Standard Deviations of Post—test according to Evaluation Styles

Group	No.	Means	standard deviations
Control Group	15	60.40	6.08
Self-evaluation	15	70.87	7.84
Peer-evaluation	15	77.40	5.19
Total	45	69.56	9.49

Table (3) shows apparent variation in means and standard deviations for the post-test because of the different categories of evaluation pattern (teacher assessment, self-assessment, and peer-assessment), and to indicate the significance statistical differences between the means, the researcher was used ANOVA according to the table (4):

Table (4)
 ANOVA for the effectiveness of Evaluation Styles in improving lingual Knowledge

Source	Sum of Squares	Degree of freedom	Mean Squares	F	Significance
Between groups	2206.178	2	1103.089	26.400	0.000
within groups	1754.933	42	41.784		
Total	3961.111	44			

Table 4 shows statistical significant differences at ($\alpha = 0.05$) due to evaluation style. To illustrate the statistical marital differences between means, Scheffe Post Hoc Test was used in Table (5):

Table (5)
 Scheffe Post Hoc Test for Evaluation Style Variable in improving lingual Knowledge

Group	Teacher-evaluation	Self-evaluation	Peer-evaluation	Means
Control Group				60.40
Self-evaluation	10.47*			70.87
Peer-evaluation	17.00*	6.53*		77.40

As seen from the table above, there are statistical significant differences at ($\alpha = 0.05$) between the evaluation of control group on the one hand and each of self-evaluation and teacher-evaluation; and the differences came in favor of both self and Peer-evaluation, it is also found that there are statistical significant differences at ($\alpha = 0.05$) between self and Peer-evaluation and differences came in favor of peer-evaluation.

Discussion, Conclusion and Recommendations

Results were discussed according to the method that highlight the most important results and provide explanations and consolidation. Depending on the theoretical literature and related studies, and highlight the most important differences and similarities with those studies, as well as recommendations and suggestions that are relevant to obtained results.

The purpose of this study is to verify the effectiveness of evaluation style in improving lingual Knowledge for Ajloun College students. The researcher discusses the results of the study according to the hypothesis as follows: First: discussing the results of the first hypothesis which stipulates that "no statistically significant differences at the level of significance ($\alpha = 0.05$) in test performance means among the students of Ajloun University College attributable to the evaluation style."

The hypothesis- test detects the rejection of this hypothesis; results indicated that there are statistical significant differences at the level of ($\alpha \leq 0.05$) between the performance means in the post-test. Post-comparisons (Scheffe Post Hoc Test) revealed that there are statistical significant differences at ($\alpha \leq 0.05$) in favor of peer-evaluation

and indicated that there is a significant difference at ($\alpha \leq 0.05$) between the performance means of two groups (the control and self-evaluation group) in favor of self-evaluation.

It is clear from the foregoing of superiority the experimental group that evaluated by peers on two groups: the control group and a group of self-evaluation; possibly due to the feedback provided by peers; it has provided students residing immediate feedback, and this is consistent with some studies like (Frankenberg, 1999; Harmer, 2006; Keh, 1990; McGarrell, 2007 and; Muncie, 2002) AlRubai (2006) confirmed the importance of feedback; since it appoints the learner to correct his wrong responses, repeats his successful responses, makes work more interesting, allows the learner to know if he is improved, affords to compete with himself, and helps him to choose right responses and work on increasing interaction, as well as the development of the mental aspects of the learner. This meets the following studies: (McGarrell, & Verbeem, 2007; Almanasreh (2006); Rollinson, 2005; Al-Hashemi (2005); Shabib (2005); Xiang, 2004; Muncie, 2002 and Hyland, 1999) that confirmed the effectiveness of feedback in correcting mistakes. Also, the present study agrees with The results of the following studies: (Al-Kharabsheh, 2005; Allen & Flippo, 2002; Nakamura, 2002; Ellition, Smith, Swanson and Stopping, 2000), and agrees also with studies that confirmed the effectiveness alternative evaluation, including: (Haajna, 2007), and the results of this study differ in terms of the superiority of the experimental group that evaluated by peers from the results of Omar study (2000), which revealed the ineffectiveness of peer evaluation, Nasr (1998) found that the peer evaluation is one of methods of little use.

The superiority of self-evaluation group to the control group revealed by post-comparisons probably due to an increase motivation created by this type of evaluation, and meets with the results of the following studies: (Al-Kharabsheh, 2005; Allen & Flippo, 2002). However, the result of this study differs from the results of a study done by (El-Koumy, 2001) which revealed the ineffectiveness of self-evaluation; Nasr 1998 has revealed that the self-evaluation is one of methods which is little use. Giving students opportunities to evaluate their own, makes them independent learners, and helps them to transfer their expertise to other similar situations; since the self-evaluation helps the learner to be more honest with himself, which increases the chance of learning by making mistakes independently; as self-group had interacted with a list of criteria of correcting test, the mistakes related to learner and not determined by others; thus self-evaluation group received is better than others because self-evaluation gives the students the opportunity to think about what they thought so evaluation is one of basic meta-cognitive skill (Mcloughlin, 2001; Marzano and Haystead, 2008; Nitko and Brookhart, 2007; Rose, 1999, and Pryor, 1998). This style gives learner the opportunity to amend procedures, and evaluation functions based on feedback from any of them; as encourage the learner to reach a high level of quality, and allows the learner to defend his arguments and evidence to justify them logically and practically (Alhilah, 2001).

Recommendations:

1. To investigate more studies about the relation between the instruction and evaluation generally and language teaching and evaluation particularly.
2. To emphasize feedback after judge to correct mistakes.

References

- Abd-AlKhaleq, M. (1996). *Language tests*. Riyadh: King Saud University.
- Abu Alam, R. (2001). *Modern theories in the measurement, evaluation and development of examination systems*, the National Center for Examinations and Educational Evaluation, Research of First Arab Conference: Examinations and Educational Evaluation, futuristic vision, p 95-120.
- Abu Amsha, K. (2011). *Linguistic evaluation. Kased center for teaching Arabic and classics*. Arab Forum: Department of Arts and Humanities.
- Abdul Hamid, J. (2002). *Trends and modern experiences in evaluating the performance of student and teacher*. Cairo: Arab Thought House.
- Abu Hashim, H. and Alsaid, M. (2010). *Future directions for psychological and educational evaluation and its applications in the field of special education*, Gulf children with special needs.
- Abu Hija, K. (2005). *Course language tests*, Yarmouk University.
- Abu Hija, S. (1989). *The impact of supplying the second graders with reading comprehension skills in developing their language abilities*. Unpublished Master's Thesis, Yarmouk University, Irbid, Jordan.
- Adel A. and Barham A. (2012). *Using Alternative Assessment Strategies in Assessing Students' Learning in Mathematics and Arabic in Jordan*. Hashemite University: Faculty of Educational sciences. <http://www.eis.hu.edu.jo/Deanshipfiles/pub103762158.pdf>
- Al-Do sari, I. (2000). *Reference Frame for educational evaluation*, 2nd edition Riyadh: Arab Education office for the Gulf States.
- Al-Faouri, O. (2006). *Building a test to measure the critical ability for outstanding students*, a worksheet, Arab Conference for gifted people: Amman.
- Alhilah, M. (2001). *Teaching methods and strategies*. Alain: University Book Dar.

- Al-kaddemat, J. (2001). *The impact of self-assessment on the writing performance for the students in the ninth grade at the 4th Directorate of Amman*, unpublished Master, University of Jordan, Jordan.
- Al-Kharabsheh, B. (2005). *The impact of the use of alternative evaluation methods on the performance of the ninth grade students in written expression*. Unpublished Master's Thesis, University of Jordan, Jordan.
- Al-Kholi, M. (2000). *Language tests*. Amman: Dar Al-Falah.
- Almanasreh, J. (2006). Standards of written expression in primary school students in Jordan. *Journal of the Education Faculty, Ain Shams University*, (13)187-123.
- Al-Hashemi, A. (2005). *Expression: its philosophy, its reality, teaching method, and corrected methods*. Amman: Dar Curriculum for Publishing and Distribution.
- Allam, S. (2000). *Measurement and Educational Evaluation: applications and modern orientations*, Cairo: Arab Thought House.
- Allam, S. (2001). *Referential diagnostic tests stake in psychological and educational fields*, Cairo: Arab Thought House.
- Allam, S. (2003). *Institutional Educational Evaluation: its principles, methodologies and applications in the school evaluation*, Cairo: Arab Thought House.
- Allen, D. & Flippo, R. (2002). Alternative assessment in preparation of literacy educators: responses from students. *Reading psychology* 23:pp.15-26.
- Allwably, A. (2003). The multi-dimensional measurement approach in the diagnosis, classification and support people with mental retardation - descriptive analytical study, *Journal of Arabian Gulf message*, volume (86) 53 - 84.
- Al-Rubaie, M. (2006b). *Supervision and Evaluation in Physical Education*. Amman: Dar Almanhaj.
- Al-Rousan, F. (1999). *Methods of measurement and diagnosis in special education*, Amman: Dar Alfekr for printing and publishing.
- Al-Sarraf, Qasim. (2002). *Measurement and Evaluation in Education*, Cairo, Dar of modern book.
- Al-Sikhan, A. (2009). *Constraints of the continuous assessment application of English language skills in the upper grades of elementary in the Qassim region from supervisors and teachers point of view*. Unpublished Master's Thesis, Faculty of Social Sciences, the Department of Education, University of Imam Muhammad bin Saud Islamic University, Riyadh.
- Altreri, A. (1997). *Psychological and educational measurement, theory, principles, applications*, Riyadh: Library of Alrushd
- Anderson. T. (2002). *Alternative assessment ideas*. University of Sioux Falls.
- Antoinette, J. (1996). Key Points of the authentic assessment portfolio. *Intervention in school & clinic*, 31, 4,252-254.
- Asp, E. (2000). *Assessment in education: Where have we been? Where we Headed?* , In Brandt, R (Ed): Education in Anew Era.
- Barziliai, A. & Frank, M. (2002). *Integration Alternative assessment in Project*. From [http: www.Leeds.ac.uk](http://www.Leeds.ac.uk).
- El-Koumy, S. (2001). *Effect of student self- assessment on Knowledge achievement and academic thinking*. Ed452731, <http://searcher.org>.
- Elliton.A, Smith. E., Swanson. I& Stopping, K. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment & Evaluation in higher education*, 25(2). Pp.146-166.
- Fadl-Allah, M. (2005). *Requirements of language evaluation under the educational standards movement*. 17th Scientific Conference, curricula and standard levels, Cairo.
- Frankenberg, G. (1999). Providing student writers with Pre- text feedback. *Elt journal*, 53(2), 100-107.
- Haajna, A. (2007). *To what extent are the Arabic language teachers use the alternative assessment strategies, and the impact of a training program on their practices of these strategies*. Unpublished PhD dissertation, Yarmouk University, Jordan.
- Harmer, J. (2006). How to teach writing. *Electronic journal of foreign language teaching*, 3(2), 242-253.
- Hyland, K. (1999). Providing productive feedback. *Etl journal*, 44(4), 279-285.
- Kawafhah, T. (2003) *Measurement, evaluation and measurement methods and diagnosis in special education*, Amman: Dar Almaseerah for publication and distribution.
- Keh, C. (1990). A model feedback in the writing process Journal and methods for implementation. *Elt* 44 (4), 294-304.
- Kelly, H. (2000). Evaluation educational software for special education. *Intervention in School & Clinic*, 36, 2,109-116.
- Mahernz, W. and Lehman E. (2003) *Measurement and Evaluation in Education and Psychology*, translated by Haitham al-Zubaidi, Alain: Dar of University Book.
- Marzano, R. J., & Haystead, M. (2008). *Making Standards Useful in the Classroom*. Alexandria: ASCD.
- McGarrell, H., & Verbeem, J. (2007). Motivating revision of drafts through for native feedback. *Elt*

- Journal*.16 (3), 228-236.
- Mcloughlin, D. (2001). *Developing Task Specific Metacognitive Skills &C Leather*; Literate Dyslexic Adults. London: Adult Dyslexia and Skills Development Centre.
- Melhem, S. (2000). *Measurement and Evaluation in Education and Psychology*, Amman: Dar Almaseerah for publication and distribution.
- Miller, M., Linn, R., & Gronlund, N. (2009). *Measurement and Assessment in Teaching* (Tenth Ed.). Columbus: Pearson.
- Ministry of Education (2005a). *Evaluation strategies and tools for Arabic language course*. Amman: Almu'alef.
- Ministry of Education (2005b). *Evaluation strategies*. Amman: Almu'alef.
- Muirhead, B. (2002). Relevant assessment for strategies online colleges & university. *Usdla journal*, 16(2) ISSN: 1537-1580.
- Muncie, J. (2002). Using writing teacher feedback in efl composing classes. *Elt journal*, 54(1), 47-53.
- Nakamura, Y. (2002). *Teacher assessment & peer assessment in practice*. Evaluation student 44 ED 4664483 from <http://searcheric.org>.
- Nasr, H. (1998). The extent of use and diversification of Arabic language teachers' methods of evaluating in the stages of general education students in Jordan. *Journal of Educational Research*, (13) 141 - 178.
- Nitko, A. J., & Brookhart, S. M. (2007). *Educational Assessment of Students* (fifth Ed.). New Jersey: Pearson Education.
- Obeidat, T. and Abu Alsamen, S. (2005). *Brain, learning and thinking*. Amman: Dar De-Bono for publication and distribution.
- Omar, Z. (2000). *Peer evaluation system for the performance of faculty members in the Division of Quran language in the Language Center at the Malaysia International Islamic University*. Unpublished PhD Dissertation, Yarmouk University, Jordan.
- Odeh, A. and Malkawi, F. (2001). *Basics of scientific research in the Education and Humanities: Elements of research, its methods and statistical analysis of data*. Irbid: library of Alkattani.
- Pryor, J. (1998). Assessment in Ghana and England: Putting reform to the test of practice. *A Journal of Comparative Education*, 28(3), 263-265.40
- Qatami, Y.; Abu-Jaber, M.; and Qatami, N. (2002). *Instruction Design*. Amman: Dar Alfekr.
- Rollinson, P. (2005). Using peer feedback in the ESL writing class. *elt Journal*, 59(10), 23-30
- Rose, M. (1999). Make Room for Rubrics. *Instructor-Intermediate*, 108(6), 30.
- Rubaie, M. (2006a). *Modern Teaching methods*. Amman: Jadara for World Book.
- Sa'adah, J. and Ibrahim, M. (2004). *The contemporary school curriculum*. Amman: Dar Alfekr.
- Sabri, M. and Rafii, M. (2001). *Educational Evaluation, its principles and procedures*, Riyadh: Library of Alrushed.
- Shabib, K. (2005). *The impact of a program based on feedback on the development of written expression skills for primary school students in Jordan*. Unpublished Ph.D. dissertation: Amman Arab University, Jordan.
- Shboul, F. (2004). *Effectiveness of using evaluation profile in teaching science in academic achievement, self-performance evaluation, and the development of social skills among the students of the seventh grade*. Unpublished PhD dissertation, Amman Arab University, Jordan.
- Taha, M. (2006). Human intelligence, contemporary trends and monetary issues. *World Series of knowledge*, V.No. 330, National Council for Culture and Arts. Kuwait: International presses.
- Topping, K. J., Smith, E. F., Swanson, I., and Elliot, A. (2000). Formative Peer Assessment of Academic Writing between Postgraduate Students. *Assessment & Evaluation in Higher Education*, 25(2), 149-170
- Wiggins. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco: Jossey- Bass Publishers.
- Xiang, W. (2004). Encouraging self-monitoring in writing by Chinese students. *Elt journal*, 58 (3), 238-246.
- Yates, R. (2000). *Rainforest reviews: an Alternative assessment*. <http://www.Multiage-education.com>.

This academic article was published by The International Institute for Science, Technology and Education (IISTE). The IISTE is a pioneer in the Open Access Publishing service based in the U.S. and Europe. The aim of the institute is Accelerating Global Knowledge Sharing.

More information about the publisher can be found in the IISTE's homepage:

<http://www.iiste.org>

CALL FOR PAPERS

The IISTE is currently hosting more than 30 peer-reviewed academic journals and collaborating with academic institutions around the world. There's no deadline for submission. **Prospective authors of IISTE journals can find the submission instruction on the following page:** <http://www.iiste.org/Journals/>

The IISTE editorial team promises to review and publish all the qualified submissions in a **fast** manner. All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Printed version of the journals is also available upon request of readers and authors.

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digital Library, NewJour, Google Scholar

