



SUMMATIVE Vs. FORMATIVE: REFLECTION ON THE SAUDI HIGHER EDUCATION ASSESSMENT SYSTEM

SABRIA S. JAWHAR

Received: 27-02-2009
Accepted: 07-05-2009

Abstract

This paper is an evaluation of the assessment system in a Saudi University for Health Sciences. The focus is mainly on looking at how the language testing system does not reflect the purpose for which it is used. It also shows how this system negatively affects the way that language education is conducted in the form of what researchers refer to as “washback”. It also suggests an alternative assessment to the present one that consists of a summative written assessment through two in-term examinations and a final one. The current type of assessment, in my opinion, fails to reflect the students’ actual learning.

Key words

Summative, formative, peer and self-assessment.

Resumen

Este artículo es una evaluación del sistema de valoración en una Universidad Saudita para Ciencias de la Salud. El principal objetivo es revisar cómo el sistema de exámenes de lengua no refleja el propósito para el cual se utiliza; también muestra cómo el sistema afecta negativamente la manera en la cual se conduce la educación de lengua en lo que los investigadores denominan “washback”. Además el artículo sugiere una valoración alterna por medio de una prueba escrita sumativa distribuida en dos exámenes parciales y uno al final del semestre. El sistema de valoración actual, a mi juicio, no refleja el aprendizaje real de los estudiantes.

Palabras clave

Valoración sumativa, formativa, auto-valoración, valoración por pares.

Introduction

In spite of continuous calls for using learner-centered approaches in language teaching, and in learning and assessment in some Saudi Higher Education institutes, the rhetoric is still far from the reality.

In most Saudi higher education institutes, teachers are still test-driven and base their teaching on that idea. They teach their students how to pass tests rather than help them to develop language skills. In addition, the students' language competence is assessed using written exams while generic and communicative skills are not tested at all (Al-Otaibi, 2004; Al-Ahaydib, 1986). Even those few teachers who tend to use new and more learner-centered approaches to teaching are faced with the traditional assessment system that does not support such relatively new ideas (c.f.2)

1. Background

In this section I will give an overview of the Saudi general education system, including teaching and testing.

1.1.1 The Saudi higher education system

Though not always immediately evident, the nature of the teaching and learning process in Saudi classrooms is influenced by the Presentation-Practice-Production (PPP) approach. Teachers seem to be particularly comfortable with this approach as it gives them control over the presented material as well as ways of assessing it (Zaid, 1993).

1.1.2 The context

I teach in a scientific university where the students are native Arabic speakers. They are approximately 18 years of age and all share the same religion, Islam, and similar cultural backgrounds. I teach English for academic purposes. My job is to prepare students to study science using the medium of the English language.

The university where I teach is a perfect example of the learning assessment techniques used in the Saudi institutes of higher education. Teachers there, as in the rest of the Kingdom, are test-driven. As with other Saudi teachers, this has had a considerable impact on teaching practices.

a. The subject knowledge

The Foundation Program in which I was involved comprises:

1. An intensive programme in English (9 hours per week).
2. An English vocabulary-building component (3 hours per week).
3. A nursing component (4 hours per week)

The intensive programme of English is supposed to be focused on enhancing the students' four language skills, with particular attention being paid to those related to communication. It is intended to bridge the gap between the students' previous experience in English and the standard required for joining the nursing programme. In other words, it lays the foundation for the students' future studies in English. It is also

aimed at increasing the students' ability to communicate using English, as a requirement for the demands of the work they will undertake. During this period students are also expected to acquire the basics for other generic skills such as problem-solving, team work and critical thinking (c.f.2.1.3).

b. Learning objectives and assessment

The university manual defines learning as:

"....learning is a dynamic process through which one acquires knowledge, understanding and skills that enable the maximizing of potential abilities. The responsibility of the students in this process is to be active in the acquisition, inquiry, internalization, and application of knowledge..... nursing is a caring process that utilizes well-defined skill in critical thinking, ethical decision-making, and independent judgment to facilitate the delivery of quality care across the life span." (Baccalaureate Nursing Program curriculum, 2006:8-9)

As one can see from the definition, generic skills such as communication and critical thinking are among the expected outcomes of the learning process in this university. Yet the present assessment system does not support the development of any of the aforementioned skills. On the contrary, it focuses on excessive summative assessment that does not help in directing students towards efficient approaches to study as will be explained in the following section.

c. Assessment techniques

At my university, lecturers use a central system of assessment that comprises of two

written, in-term examinations, and one final examination. These exams are made up mainly of questions on discrete grammar points, vocabulary items and very short compositions. As such, students are not evaluated in terms of their ability to use the language for communication. Rather, they are evaluated according to how much they know about it.

The cut-off point is 60% and the result is either pass or fail, which makes it an accountability test of sorts. The course is non-credit awarding however. If a student fails the Foundation Program, she will be given the chance to repeat it once. If she fails again, she will be dismissed.

In general, teachers and administrators treat assessment as the ultimate goal of the teaching and learning process. Previous research suggests that this enhances the testing culture in the students' minds and limits their learning development (Ames, 1992; Torrance and Pryor, 1998). The assessment role in shaping the students' learning and development is overlooked.

2. Forms of assessment

In this section I will shed light on some of the assessment-related terminology, especially those that I will be referring to in this paper. At the same time, I will discuss some of the pros and cons of using these assessment methods and connect them to my context where possible.

2.1 Summative assessment

Researchers define summative assessment (hereafter SA) as a tool that is used to record

the overall achievement of the learner at a specific time. Clarke (1998), for instance, describes SA as 'snapshot testing', which reflects what a child can and cannot do at a given time. While the same can be said about adults, Gipps and Murphy (1994) believe that it is a tool that provides evaluative information by the use of a task. This information, they state, can be assertively used to make a comparison among students or schools. SA is mainly used for validation and accreditation. This limited use, however, has resulted in a negative connotation that is associated with the term socially and within educational institutes.

Educators, especially in higher education, are exerting efforts to distance themselves, at least literally, from the use of the term. However, despite the negative light in which educators see summative assessment, Earl (2003) states that SA is still the dominant type of assessment in schools. He believes that there are going to be always "...milestones where summative assessment is required." (Earl, 2003: p.22)

To Earl (ibid.), classroom assessment has a multitude of purposes, many of which are contradictory. For example, 'traditional' classroom assessment and reporting processes facilitate comparison amongst students, fulfill accountability demands but, paradoxically, do not provide any specific information about what the students have achieved.

Another problem that has been associated with the broad use of this kind of assessment in my context is its inability to assess other aspects of students learning, as identified

in the university learning objectives (c.f.2.1.3). In my context, SA is used as the only tool by which judgments are made concerning the students' level of achievement. This sole use of SA in my university has deprived the students any sort of constructive contextualized feedback or direction during their learning (Koller, 2005; Allal and Mottiez-Lopez, 2005).

2.2 Formative assessment

Formative assessment (hereafter FA) has emerged as a reaction to the need to suppress the negative effects of summative assessment. However, there is still some degree of confusion and debate concerning what formative assessment is and why it should be used in the first place (Clarke 2001; Taras 2008). Many teachers however, especially in my context, think that it is an additional tool that they can use inside the classroom throughout the semester, but one that has nothing to do with the students' learning or grading.

There is currently a wave of support for formative assessment, with the phrase 'assessment for learning' being adopted to highlight the role of assessment in the teaching and learning process (e.g. <http://www.assessment-reform-group.org>). However, it is crucial to note that the current 'assessment for learning' term has come into existence as a result of educational researchers' increasing criticism of traditional assessment and grading practices.

Black and William (1998) questioned the ability of formative assessment to bring about an improvement in learning. They

carried out an extensive survey of the available research literature over a period of nine years. The aim of their study was to find quantitative evidence to support the relationship between formative assessment and learning. They confirmed the strong relationship between formative assessment and efficient pedagogy. However, although they failed to show the impact of the context of the application in such studies, their research is considered as a comprehensive framework for FA.

Proponents of FA base their arguments on the constructivist views of learning (Vygotsky 1978). Sedler (1989) for instance, emphasized the importance of the students understanding the desired goals and how far or how close they are to achieving them. To him, learning should include the engagement of students in activities that help them to bridge the gap between their current level and the desired one.

Based on the previous discussion, one can conclude that, unlike SA, FA is a continuous process that is embedded as part of the teaching routine. When carried out as a daily practice, it informs the teachers about the effectiveness of their teaching methods as well as the progress of their students' learning. Some might argue that SA can play the same role (Black and Wiliam 2003; Black et al. 2003, 2004). The answer is that this is possible but the difference is that FA provides information when it is still possible for the teachers and the students to make changes.

Another advantage of FA is also its ability to overcome the negative effect of social accountability or what Messick (1985) refers

to as 'social consequences', that have been associated with SA. It works like a guideline that helps the teacher to know in advance what the next step is, and allows him/her to make decisions accordingly. It is also marked by the importance it gives to formative feedback and to students' involvement, as crucial practices for effective implementation. This involvement can take different forms, depending on the kind of task the teacher is implementing. However, it can be categorised under self-, peer- and co-assessment (c.f. 5).

Formative assessment, nevertheless, requires more time and effort on the teacher's part. It also requires training for both teachers and students. In general, this kind of assessment is absent from the Saudi classroom for different reasons including the central assessment system that limits the teachers to the use of written exams (Al-Otaibi, 2004; Al-Ahaydib, 1986). The same system also affects the students' and the teachers' understanding of learning and teaching (i.e. negative washback c.f.3.3.).

The previous discussion suggests that the principles of formative assessment such as constructive feedback, students' needs and involvement are promising alternatives to the currently used method of assessment in the Saudi context (Ames, 1992; Blue, 1988; Lee and Suini 2005; Falchikov, 2005).

2.3 Validity and Washback

Validity is a term that is used to refer to the extent to which the test actually measures what it claims to measure (Brown, 1996, p. 231). It is also used to refer to the appropriateness of the inferences and

decisions made on the basis of a test's results. Sireci (2007) reviewed the available literature with regard to test validity and came up with a list of conclusions among which he said that validity "...is not a property of the test. Rather, it refers to the use of the test for a particular purpose," (p.477).

Messick (1989) believes that one of the main elements that put any test's validity at high risk is construct under-representation. According to him, this under-representation occurs when the test is constructed in such a way that does not include "...important dimensions of facets of the construct" (p.34) which is something that is apparent with regard to the assessment system used in my university.

Washback, on the other hand, refers to the effect that the use of a particular test has on the teaching and learning process and that would not have happened if the test were not used (Messick:1985). Recently, the term ' washback' has received a good deal of attention to the extent that some researchers (Messick, 1996; Pintrich, Wolters & Baxter, 2000) claim that a test's validity can only be determined by the extent to which it has a positive or negative washback effect on learning and teaching . Messick (1985) also believes that washback is "...an instance of the consequential aspect of construct validity" (p.241).

In the light of the above discussion, the problem with the assessment system in the Saudi higher education context can be summarized as follows:

1. The written tests used focus on only discrete grammatical points and two of

the four language skills, namely writing and reading. At the same time, it does not assess listening and speaking. This fact makes its construct validity subject to questioning (Messik 1989).

2. The overuse of summative assessment does not precisely measure all the learning outcomes or help students to develop, due to the absence of constructive feedback (Lee, 2003; Taras, 2008).
3. The relationship between assessment on the one hand and teaching and learning on the other, is not taken into consideration, an aspect that can result in a sort of negative washback (Black and William 1998a; Torrance and Pryor, 1998; Biggs, 1999).
4. Though generic skills (e.g. problem solving, decision making and team work) are considered the ultimate goal of the learning process, the assessment method does not reflect this.

3. Rationale

Despite the importance of Item Analysis in a language test's evaluation, this paper is limited to evaluating the adequacy of the assessment system used in Saudi higher education, in comparison to the purpose for which it is used, as well as the desired learning outcomes. It also looks at the impact of the use of such an assessment system on teaching practices.

Kane (2006) notes that it is impossible to prove that a test is valid nor even to prove that we are measuring what we think we are

measuring. Sireci (2007), on the other hand, stated that any conceptualization of validity theory must recognize that what is to be validated is not the test itself, but the use of the test for a particular purpose. He believes "...that validity can be evaluated only with respect to a specific testing purpose" (p.478). He also said that "Unintended negative consequences might not point to problems in tests per se but may indicate problems within a testing system" (p.480)

I believe that the problem in my context is not related to the test itself but more to the use of the test for purposes that go beyond what it is intended to assess. The following section will discuss some of the available alternative to the sole use of summative written examinations.

4. Suggested alternatives to the present mode of assessment

Recent years have witnessed a paradigm shift in teaching and learning practices, from the teacher- to the learner-centred classroom. It has also been accompanied by an increasing emphasis on the learner's role, not only in the learning process, but also in the assessment practice with special attention being given to the relationship between the two (Driscoll and Wood, 2007; Cheng & Warren, 2005; Bresciani et al. 2004; Luoma and Tarnanen, 2003).

4.1 Redefining assessment

Before suggesting any alternative, I believe that the concept of assessment used in the Saudi context should be revisited. A good definition of assessment is found on the University of Oregon website. Here,

assessment is defined as:

"The process of gathering and discussing information from multiple and diverse sources in order to develop a deep understanding of what students know, understand, and can do with their knowledge as a result of their educational experiences; the process culminates when assessment results are used to improve subsequent learning,"<http://tep.uoregon.edu/workshops/teachertraining/learnercentered/assessing/assessing.html>

As the definition indicates, assessment should not be an end in itself and it should not be reduced merely to summative written exams at the end of the course. On the contrary, it should be a tool for gathering information about the students' learning outcomes. It should also be a holistic process that includes strategic planning as well as continuous formative feedback from different sources.

4.2 Assessment and learning

The literature shows that the link between assessment and the learning process tends to be overlooked by many teachers and researchers (Torrance and Pryor, 1998). In many Eastern and Middle Eastern countries, such as Saudi Arabia, assessment is still perceived as the last step of the curriculum development process (c.f.2). This misperception of assessment has had its effect on those students who are denied the opportunity to become involved as active participants in the whole process.

William (2003) draws a distinction between formative assessment and assessment for learning, conceptualizing assessment for

learning as the 'purpose', and formative assessment as the 'process'.

Discussing the concept of assessment for learning, Blue (1988) believes that learning enhancement is associated with the learners' involvement in assessment, as well as needs analysis. He emphasizes the importance of such an inclusion, not only for the students' learning outcomes, but also in order to minimize the teachers' workload. It is worth noting here the fact that this exclusion is not a result of bad intentions as much as ignorance of the multi purposes of assessment from the teachers' point of view. Torrance and Pryor (1998) explored this issue in a paper that aimed to investigate teachers' understanding of the term 'assessment'. They found that teachers perceive assessment as an activity discrete from actual teaching, so they used it as a tool for gathering information for a 'third party', whether school or parents, a situation that is similar to that in my context. They stated that teachers did not see the benefit of assessment, either for themselves or for their students. Moreover, teachers were of the opinion that assessment has a negative effect on the students' motivation and learning.

Ames's (1992) stand supports that of Torrance and Pryor (op.cit). She believes that the way by which the students are evaluated can affect their motivation but, to her, the key issue is the students' perception of the meaning of evaluation. She claims that the way that evaluation is structured can affect the orientation of students toward different goals. This explains my students' test-driven orientation towards knowledge. She also warns against the negative effect of social comparisons that might operate

among students, especially when it comes to winning or out-performing their peers, a thing that happens frequently in SA. She states that such comparisons affect the students' perception of their ability, or lack of it. Students' use of effective learning strategies and problem-solving strategies depends on the appreciation of the effort they put into the task. Grades, she adds, can have a harmful effect on the efforts to implement more learning strategies because they support social comparisons. However, she states, they can still be useful, but only if they are used alongside feedback that serves to improve the work. Under these conditions, the students are more likely to use such strategies.

Harlen and Deakin Crick (2003) downgraded the importance given to the relationship between testing and students' motivation. They said that only a few studies offer evidence that supports this relationship. However, they called for further research in this respect.

Regardless of the effect of assessment on the students' motivation, the previous review suggests that constructive feedback included in the formative assessment is an important strategy to help students develop their learning. It gives the students an insight into what they are doing well. It also provides the students with precise input with regard to how to progress in their learning. Students in my context lack this kind of feedback. This can be obtained from different sources among which are the following:

4.2.1 On-line assessment

Using especially designed computer

programs for assessment is one of the potential alternatives to traditional assessment for different reasons, including providing immediate feedback and offering learning opportunities. It gives the students more flexibility when it comes to time and setting. It also minimizes teacher workload (Brown, 1987; Kearsleg, 2003; Johnson & Aragon, 2003). However, in my context, this is not applicable for several reasons including the availability of staff who are able to deal with computer programs and who are able to guide the students throughout the semester. We have also to consider the availability of equal access to computers as well as computer literacy among students. Not all of my students are good when it comes to computer use. Some of them they do not even have access to computers. Assessment should not be a subjective process that serves only one group. Chances of plagiarism should be minimized while opportunities for self-learning should be explored (Heberling, 2002).

4.2.2 Self- and peer-assessment

These are among the most highly recommended approaches to assessment, especially at higher education level (Brud et al., 2003; Elliot & Higgins, 2005). Self-assessment is believed to help students to develop generic and higher intellectual skills such as problem-solving and critical thinking. It is also argued that it helps to minimize teacher workload. It promotes learner autonomy and helps the teacher to meet the students' expectations and to be aware of their opinions (Dochy et al., 1999; Taras, 2001; Wilson & Fowler, 2005). However, students should be trained to carry out self-assessment in order to overcome the

problem of validity and credibility (Ross, 2006).

Taras (2001) tackled this problem and suggested three steps towards better self-assessment. She believes that the learners should start by getting grades for their self-assessment exercises. Teacher guidance should be available to the students, especially at the initial stages, as this will give them a sense of security.

Taras suggested offering teacher's feedback as a second step. She believes that it would help the students to recognize their mistakes and learn from them. Finally, she proposes imposing the completion of the self-assessment exercises as a condition for receiving grades.

Peer assessment and reflection practices are also believed to positively influence the students' learning approaches, to help solve the problem of large classes, to allow students to reflect on their learning and to become more independent (Wilson and Fowler, 2005; Dochy et al., 1999; Strachan and Wilcox, 1996). Moreover, it is believed to encourage lower level students to move towards a deeper learning and to develop "...independent research, collaboration and communication" (Strachan and Wilcox 1996. p.343). By introducing self and peer-assessment, however, educators should be aware of the students' backgrounds and their attitudes towards such alternatives (Winnie and Warren, 2005; Lee and Suini, 2005; Biggs, 1999).

Littlewood (1999) indicated that East Asian students expected the teacher, as the holder of authority and knowledge, to be responsible

for the assessment of learning. Similarly, Saudi students expect the teacher to be the one responsible for assessment. This fact might influence their attitude towards such alternatives, and call for pre-assessment training sessions and attitude surveys.

5. Future implications of assessment in the Saudi context

From what has been mentioned we come to the conclusion that assessment is not an end in itself. On the contrary, it is a crucial part of the curriculum and should be integrated into the teaching and learning process. If given enough attention and carefully designed, assessment can effectively influence student learning.

It is true that lecturers at some university, like mine, use a central assessment policy that reduces the assessment to written summative exams, yet there is always room for change and improvement. However, it is worth mentioning the fact that change is not an easy process that can be done overnight, especially when it comes to educational practices. It should be introduced gradually, but as part of a holistic and strategic plan. With regard to my context, several changes should take place at both the policy and the test-designing levels. But until teachers and students are prepared to move from the current paper and pen examinations to an assessment system that takes advantage of advancements in technology (Raikes & Harding, 2003), the following changes should be considered:

1. Since it is not easy to move from one end of the spectrum to the other without causing confusion on the part of both

teachers and students, special attention should be given as to how to implement summative assessment in such a way that makes it possible to provide the required information that matches the purposes for which it is used. To do so, we have to carry out an academic needs analysis to find out the language skills that our students will need and include them in the taught materials. Following this, we have to make sure that the assessment reflects these introduced materials. In my case, listening and speaking should be included as part of the assessment process. Also the other generic and cognitive skills such as teamwork and problem-solving should be covered equally in the introduced and the assessed material.

2. The item specifications used in the written exams should also be developed in accordance with the students' needs, during and after their years of studying. Within those specifications, qualities such as validity, reliability and practicality should also be closely considered.
3. The teacher should take advantage of the fact that students study what they expect to be assessed on, and focus their teaching practices around specific learning outcomes and relate them to assessment. The 100 score that is currently assigned to three written examinations given in a specific and limited time, should be redistributed to include other activities such as group work, peer- and self-assessment.
4. The teachers should give more importance to providing the students with

more constructive feedback following every task or examination, so that they know how well they have performed and what should they do to improve in the future.

5. Teachers and students should be trained to use alternative assessment techniques such as peer- and self-assessment after their attention is drawn to the advantages of such techniques for both the learning and teaching. These techniques will help students to develop their cognitive and generic skills.
6. A balance should be kept between formative assessment that measures the students' learning outcomes and development and summative assessment upon which the students' level will be judged.
7. Students should be involved in drawing up the assessment criteria. At the same time, those criteria should be made transparent and detailed. To this end, I suggest the use of a detailed rubric sheet. This is not to pin down everything for them, but to help them to know what is expected of them, and to increase the effectiveness of their learning.

6. Conclusion

In this paper I have briefly addressed the importance of assessment, not as an end in itself, but as a contribution to the students' learning development. Citing my university as an example, the paper evaluates the learning assessment system in Saudi higher education institutes and suggests changes in the assessment policy. It also suggests

the use of peer- and self-assessment as alternative to the current one. This change, however, should be gradual and proceed as part of an intensive review of the whole curriculum, including institutional policy and course goals.

Finally, since this paper is based on my own experience, it is important to understand that we cannot generalize from it. Rather, it should be seen as an eye-opener that could serve as a springboard for further large scale and empirical research into the area of assessment.

References

Al-Ahaydib, M. (1996). 'Teaching English as a foreign language in the intermediate and secondary schools of Saudi Arabia: A diagnostic study.' Unpublished doctoral thesis, University of Kansas.

Allal, L., and Motiez-Lopez, L. (2005). Formative assessment of learning: A review of publications in French. In Centre for Educational Research and Innovation (2005). (CERI). *Formative assessment: Improving learning in secondary classrooms* (pp 241–264). Paris: OECD.

Al-Qahtani, F. (2003). 'Introducing the target culture to develop the sociolinguistic competence of EFL students: Views and attitudes of Saudi EFL teachers at selected male middle schools in Riyadh, Saudi Arabia. Dissertation Abstracts International, 64 (07), 2412. (UMI No. AAT 3097634).

Ames, C. (1992) 'Classroom: goals, structures and student motivation', *Journal of Educational Psychology*, 84, pp. 261-271.

Biggs, J. 1999, *Teaching for quality learning*

at university, Buckingham, UK: SRHE & Open University Press.

Black, P. and William, D. (1998) *Inside the black box: Raising Standards through Classroom Assessment*. London: Kings College.

Black, P., & Wiliam, D. (2003). 'In praise of educational research': Formative assessment. *British Educational Research Journal*, 29(5), 623–637.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead: Open University Press.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. *Phi Delta Kappan*, September 2004, 9–21.

Blue, G. M. (1988), 'Self-assessment: the limits of learner independence' in Brookes, A and P Grundy (Eds) 1988. *Individualization and autonomy in language learning*, (pp100-118). Modern English Publications in association with The British Council.

Bresciani, M.J., Zelna, C.L., and Anderson, J.A. 2004, 'Assessing Student Learning and Development', National Association of Student Personnel Administrators.

Brown, H. (1987). *Principles of language learning and teaching*. Englewood Cliffs, New Jersey: Prentice Hall Regents.

Brown, J. D. (1996). *Testing in language programs*. Upper Saddle River, NJ: Prentice Hall Regents.

Burd, E., Drummond, S., Hodgson, B. (2003) *Using self and peer assessment for group work*. Available from: http://www.ics.ltsn.ac.uk/pub/conf2003/using_peer_and_self.htm

Clarke, S. (1998) *Targetting Assessment in Primary Classroom*. London: Hodder and Stoughton.

Clarke, S. (2001) *Unblocking Formative Assessment- Practical strategies for Enhancing Pupils' Learning in the Primary Classroom* London: Hodder and Stoughton.

Dochy, F., M. Segers and D. Sluijsmans. (1999), 'The use of self-, peer and co-assessment in higher education: A review', *Studies in Higher Education*, 24(3), pp. 331-350.

Earl, L. M. (2003) *Assessment as Learning: using classroom assessment to maximize student learning*. London: Sage.

Elliott, N. and Higgins, A. (2005) Self and peer assessment – does it make a difference to student group work? *Nurse Education in Practice* (5) p.40-48.

Falchikov, N. (2005), *Improving Assessment through Student Involvement*. New York: Routledge Falmer.

Gipps, C. and Murphy, P. (1994) *A Fair Test? Assessment, achievement and equity*. Buckingham: Open University Press.

Harlen, W. and Crick, R. D. (2003) 'Testing the Motivation for Learning', *assessment in Education: Principles and Practice*, 10, (2).

Heberling, M. (2002). Maintaining academic integrity in on-line education. *Online*

Journal of Distance Learning Administration, 5(2). [Online] Available at: <http://www.westga.edu/~distance/ojdla/spring51/heberling51.html>

Johnson, S. & Aragon, S. (2003). An instructional strategy framework for online learning environments. *New Directions for Adult*

and Continuing Education, 100, 31-42.

Kane, M. T. (2006) 'Validation', in Brennan, R. L.(ed), *Educational measurement.4* Westport, CT: American Council on Education/ Praeger, pp. 17-64.

Kearsley, G. (2000). *Online Education: Learning and Teaching in Cyberspace*. Toronto: Nelson Thomson Learning.

King Saud bin Abdulaziz University for Health Sciences, (2006), Baccalaureate Nursing Program Curriculum. Unpublished manual.

Koller, O. (2005). Formative assessment in classrooms: A review of the empirical German literature. In Centre for Educational Research and Innovation (2005) (CERI) Formative Assessment: Improving Learning in Secondary Classrooms (pp 265–275). Paris: OECD.

Lee Sang-Ki, and C. Suini.(2005), 'Learner Involvement in Self- and Peer- Assessment of Task-Based Oral Performance', *Language Research*, 41(3), 711-35.

Littlewood, W. (1999), 'Defining and developing autonomy in East Asian contexts', *.Applied Linguistics*, 20(1), 71-94.

Liu, K P et al. (2003), 'Approaches to improve Chinese learners' English writing', *Foreign Language Education*, 24(6), 68-71.

Luoma, S. and M. Tarnanen (2003), 'Creating a self-rating instrument for second language writing: from idea to implementation', *Language Testing*, 20(4), 440-465.

Messick , S. (1996) Validity and washback in language testing, *Language Testing*, 13(3), 241-256.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13-104).

New York: Macmillan.

Patri, M. (2002), 'The influence of peer feedback on self- and peer-assessment of oral skills', *Language Testing*, 19(2), 109-131.

Raikes, N. And Harding, R.(2003), The horseless carriage stage: replacing conventional measures, *Assessment in Education*,10(3), 267-77.

Ross, J. A. (2006), The reliability, validity and utility of self-assessment, *Practical assessment Research and Evaluation*, 11(10), 1-13. Available online at: <http://pareonline.net/pdf/v11n10.pdf>

Sadler, D. R.(1989), Formative assessment and the design of instructional system *.Instructional Science*,18(2), 119-144.

Sireci, S. G. (2007) 'On Validity Theory and Test Validation', *Educational Researcher*, 36, pp. 477.

Strachan, I.B. and S. Wilcox (1996), 'Peer and self assessment of group work:

Developing an effective response to increased enrolment in a third-year course in microclimatology', *Journal of Geography in Higher Education*, 20 (3), 343-353.

Taras, M. (2008), 'Summative and formative assessment: Perceptions and realities', *Active Learning in Higher Education*, 9(2), 172-192.

Taras, M. (2001), 'The Use of Tutor Feedback and Student Self-assessment in Summative Assessment Tasks: towards transparency for students and for tutors', *Assessment and Evaluation in Higher education*, 26 (6), 605-614.

Torrance, H. and Pryor, J. (1998) *Investigating Formative Assessment: Teaching and Learning in the Classroom*. Guildford: Biddles Ltd.

University of Oregon, Teacher training, <http://tep.uoregon.edu/workshops/teachertraining/learnercentered/assessing/defintion.html> (retrieved on 15 September 2008)

Vygotsky, L. (1978). Interaction between Learning and Development (pp. 79-91). In *Mind in Society*. (Trans. M. Cole). Cambridge, MA: Harvard University Press.

Winnie, C. and M. Warren (2005), 'Peer Assessment of Language Proficiency', *Language Testing*, 22(1). 93-121.

Wilson, K. and J. Fowler (2005), 'Assessing the impact of learning environments on students' approaches to learning: comparing conventional and action learning designs', *Assessment and Evaluation in Higher Education*. 30 (1), pp.87-101.

Zaid, M. (1993), *Analysis of the current system of teaching English as a foreign language In the Saudi Arabian intermediate schools*. Unpublished doctoral dissertation, University of Colorado, Colorado.

THE AUTHOR

SABRIA S. JAWHAR is an PhD. student at Newcastle upon Tyne University, UK. She is also a lecturer at King Saudi Bin Abdul Aziz University for Health Sciences, Nursing College, Jeddah, KSA.

Contact: sabria.jawhar@yahoo.com