THE CHANGING PURPOSES AND PRINCIPLES OF ASSESSMENT IN SAUDI ARABIA: IMPLICATIONS FOR TEACHERS

Abeer Sultan Althaqafi[i]
Assistant Professor
English Language Institute (ELI),
King Abdulaziz University (KAU),
Jeddah, Saudi Arabia

Abstract:

Assessment represents the core of learning attainment. It is considered the most contentious area of education both historically and in contemporary debate. In addition, it is a major educational issue where rapid changes are taking place. It plays a crucial role since it provides teachers and learners with the needed information in which they can apply many decisions. Such decisions can affect pupils' attitudes and choices in life. Therefore, it is of paramount importance to identify the purpose of assessment and make it beneficial to the pupils. The aim of this paper is to discuss the changing trends and purposes of assessment from the 60s to present day. In this paper, we will try to answer some of the following questions: Why do we need to assess? What are the main purposes of assessment? And how should we assess? In the following paragraphs, we will shed light on the traditional view of assessment together with the new and old trends in this regard; also, we will investigate the changing purposes of assessment in Saudi Arabia. In addition, there will be a brief discussion of the emerging trends in this field as well as mentioning some of the implications for teaching and learning.

Keywords: Evaluation, testing, examination, assessment, summative assessment, formative assessment, and computer assisted assessment.

1. The Traditional/Contemporary Trends in Assessment

First of all, we need to be clear about what is meant by assessment in the field of education, and what the differences are between evaluation, assessment, testing and examinations. 'Assessment' is a general term encompassing measurement instruments concerned with how well the student has done, conducted by a teacher, an examiner or even the pupil himself/herself (Jones and Bray, 1986). 'Evaluation' cannot take place

---

[i] Correspondence: email abalthagafi@kau.edu.sa

without assessment; it judges whether it was worth doing in the first place. A 'test' is a particular measurement set up for the purpose of making an assessment, whilst an 'examination' is a combination of several tests and other assessment procedures (Ibid., 1986). In fact, it can be found that assessment is equated with tests and examinations.

A traditional notion of 'assessment' in most schools and colleges in Saudi Arabia is regarded as 'a vision of tests or examinations, certificate and lists of marks.' This was considered as reliable precise measures of achievement since the role of the assessment was mainly to provide us with outcomes of the performance of others. The assessment was carried out in a very formal restricted atmosphere under controlled conditions. Some major assessments were undertaken under the regulations of national examination boards.

Unfortunately, assessment in the past was mainly academically oriented and highly dependent on memory. There has been a huge emphasis in examinations on only the productivity. Nevertheless, in the last two decades, the concept of assessment has changed from the traditional notion of 'testing' for selected purposes, to a much wider concept of multiple purposes integrated with the curriculum. In addition, there is a strong desire now to improve students' motivation and involvement through engaging them in recording their own levels of achievements and in setting their own goals. In short, the following table, based on my reading of the literature conducted by Jones and Bray (1986), summarises some of the major changes in purposes and trends of assessment in the past and present day.

Table 1: Changing trends and purposes of assessment (Adapted from Jones and Bray, 1986)

Former Trends	Recent Trends
Summative assessment in a formal setting.	Formative and informal assessment.
Infrequent examinations on special occasions.	Continuous assessment, usually as a normal part of the teaching/learning process.
Comparing pupils with each other in order to indicate final achievement and ranking.	Comparing pupils' achievements to provide feedback and improve performance.
Assessment in artificial situations (exams halls) dominated by the notion to be cheat proof.	Assessment in situations more akin to real life, e.g. 'open book' examinations, practical projects.
External responsibility for setting and marking examinations.	Internal responsibility for assessment; professionalism of teachers; self-assessment by pupils.
Recall of course content.	Emphasis on learning process, e.g. information retrieval, study skills.
Assessment is concerned with academic achievement and content within established subjects.	Recognition given to skills such as listening, speaking, attitudes, practical skills, personal and social development.

Less use of technology in assessment.	Computer-assisted assessment.
Assessment as a self-continued activity.	Assessment seen as an integral part of the curriculum and of teaching/learning.

The above table summarises old and new trends in the system of assessment, and from the overall view we can see that there were some disadvantages to such an antiquated approach: it did not test understanding or application of the new input; it did not reflect on what happens in real life; it did not appear to test persistent effort or long-term memory over a long period of time; and it did encourage critical thinking. However, it should be highlighted here that some schools and teachers still utilise the traditional old forms of assessment as in most developing countries. However, most schools are probably half-way between the two trends whilst others have jumped to the new vision of assessment. Assessment is clearly an ongoing changing aspect, so in the following paragraph we will try to find out why does it change?

1.1 The Need for Change

As we can see from the above discussion, assessment is changing and will continue doing so due to the fact that students are facing a world which demands new knowledge and abilities. In fact, assessment is changing for many reasons: firstly, the shift of our societies from the industrial age, where a person could survive with basic reading, writing and standard arithmetic skills, to an information age, which demands the ability to access, interpret, analyse and make crucial decisions. This 'knowledge explosion' together with the increase in the amount of available facts, have placed much greater emphasis on learning process and assessment (Bray, 1986; National Center on Education and the Economy, 1989; and U.S. Department of Labor, 1991; Bond, 1992). Moreover, according to McNamara (2001), there is a growing awareness of the fundamentally social role of assessment challenges us to rethink our priorities. Such awareness has been brought about by influential work on validity and by the intellectual changes by postmodernism.

Furthermore, it is very important when speaking about the change to clarify those assessment decisions of making a change should be related to the purpose of assessment and the content to be assessed. In fact, teachers as well as learners are required to be involved in the changes and to incorporate the changes into their practice. It is believed that when changing assessment without having input from the two poles of the learning process (teachers and students), results often become in resistance or ineffective to change (Corbett and Wilson, 1991; Borg and Edmett, 2018). We as teachers should not ignore the goal of selecting the measurement of assessment that both match the outcomes to be assessed and the purpose of the assessment.

In order to set an example for such ongoing changes, the following section will outline the changing purposes of assessment and examinations focusing on Saudi Arabia.

2. Changing Purposes and Principles: an Extract from National Curriculum in Saudi Arabia

From my own experience as a former Saudi student and a teacher in a number of schools and institutions besides the knowledge of the topic, it can be said that the former situation of assessment procedures in most schools and colleges in the Kingdom of Saudi Arabia, was carried out towards the latter stages of a course or unit of work, undertaken under strictly controlled conditions solely for the purpose of assessing students' performance. The results were given to the pupils in a coded language of grades which provided them with limited insight of their performance. Moreover, there was a widespread practice of not returning test papers to pupils as well as not revealing procedures of marking or even allowing students to engage in any kind of post-assessment discussion. All the examinations were undertaken under the auspices of Saudi Ministry of Education. Furthermore, it should be mentioned that the assessment was mainly subject-based and this had many limitations such as those suggested by Horton (1990), for example, additional constraints on the teacher inside the classroom and limited learning opportunities.

In fact, Saudi Arabia is like other countries that have embarked on substantial reform in education, tracing the broadening notion of assessment for the good of pupils. Therefore, the Saudi National Examination Board in the 1980s, based on the information derived from assessment in most schools and universities, has set certain measures for the assessment to be capable of serving several purposes. My review of the Saudi exam documents suggests that the following types of assessment are used, (there are no documents relating to the Saudi assessment system included, since most of them are in Arabic). The following diagram will show these emerged trends:

**Figure 3.1**: The purposes of assessment in the 1980s
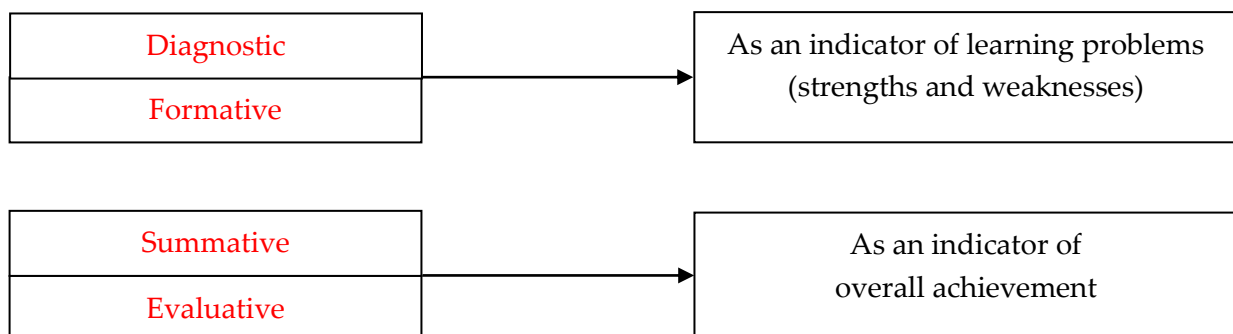
This diagram shows that formative assessment was and (still) used for grading, ranking and selecting purposes. It is designed to help pupils by identifying positive or negative outcomes that shall be recognised and discussed with each designated pupil or pupils and then plan for the appropriate followed steps. On the other hand, I should make some clarification here, as you can see that there is no clear boundary between formative and diagnostic purposes. If the assessment is designed in formative terms, it is also likely to supply some diagnostic information of strengths and weaknesses. The formative diagnostic type of assessment is of great value as indicators of learning problems that require further attention. Therefore, it is clearly seen that the situation of assessing at most Saudi schools and at the first years at most universities and colleges is mainly formative. In fact, assessment procedures designed for formative purposes could meet all the requirements of national assessment at ages before eighteen. At age eighteen, when the pupil is required to undertake a terminal examination at the end of the compulsory period of education, the focus shifts from formative to summative. This application of this terminal examination serves a selection of functions of assessment (rewarding pupils who seem to do it best) which was a dominant influence on most comprehensive schools, regarding the way in which the curriculum is defined as well as in terms of the images of success and failure that are associated with the curriculum.

In addition, the Saudi Ministry of Education makes significant efforts to have a clear summative judgement or aggregation of the data of assessment for the purpose of recording pupils' overall achievement in a systematic manner and then provide relevant information for other educational service to be assessed and reported upon.

Since the 1990s, the notion of assessment in Saudi Arabia has been changing as is the case for a large number of countries (Ajyaal, 2008). It was broadened to cover several functions including: the motivation of students to learn, improving teaching, and enabling learners and others to make rational choices about courses, careers and others activities. According to the literature on assessment produced by Barry O'Sullivan who is very active in language testing, and currently works with government ministers, universities and test developers in the Middle East and in a number of other countries, the following types of assessment are used in most Gulf countries including Saudi Arabia. The following figure will show some of the major objectives of assessment in the last two decades:

Figure 3.2: The new purposes of assessment in the 1990s
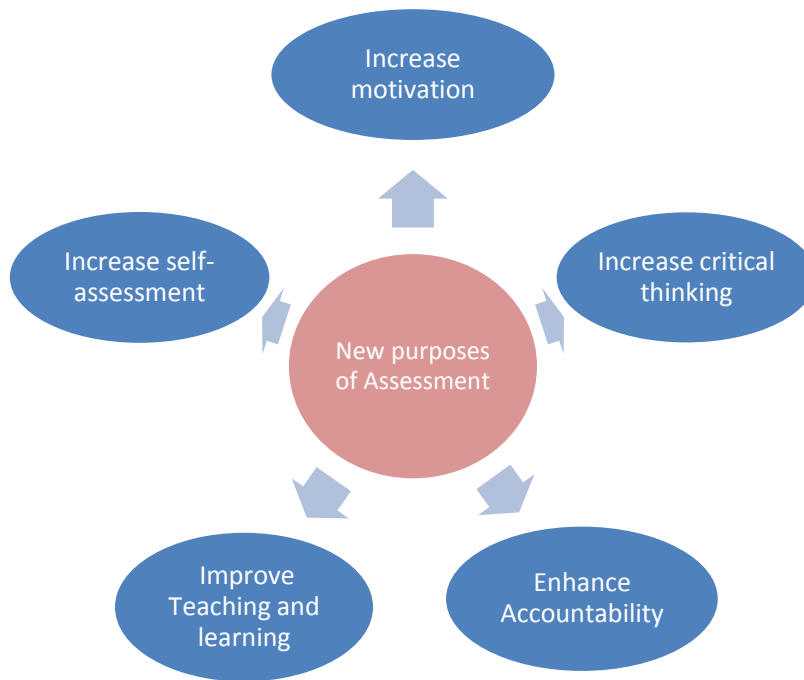
This diagram shows that the notion of assessment has been broadened to serve other purposes which have the potential to help students to obtain a better understanding. According to many researchers, such as Newstead (2004), this new insight to assessment helps to assess students' ability to integrate, transform and use knowledge purposefully since they have to develop their own understanding and critical thinking by transforming and reorganising established knowledge and procedures. Thanks to the new trends in assessment, Saudi students became more critical analysts of their own work since teachers began to incorporate self-assessment into their programmes.

However, the new notion of assessment and educational achievement is in the midst of a full-blown national debate which has been fuelled by the National Examination Boards proposals about the nature of curricula, and whether the assessment should concentrate on basic skills of the core areas of the curriculum whilst the other areas can be taught but do not require such formalised assessment. Also, at the heart of the debate is how far to continue to change the assessment system according to the nature and purpose of education. Educationalists believe that to choose between different approaches of assessment means to choose between different curriculum emphases and consequently there will be new implications for teaching and learning strategies.

From a Saudi perspective, it should be mentioned that the current approaches of assessment and the various purposes it aims to serve, have resulted in the whole situation becoming a political battlefield. The emphasis given to each of these purposes is influenced to a high extent by the social and political context in which assessment takes place. The educational policy in Saudi Arabia has become increasingly driven by

the need to measure the outcomes of its individuals and report against national standards to account for public expenditure. This is the government's main aim in the assessment of the 'Educational Health' of the country. In fact, in Saudi Arabia as well as in any other countries the need still arises for a more flexible measurement focusing on improving learners' opportunities to obtain a good attainment through the learning process as well as improving capabilities rather than judging them (Ajyaal, 2008).

To sum up, it appears from the above discussion that we can summarise the general situation of assessment in Saudi Arabia as follows:

- Assessment system in Saudi Arabia is an area where rapid changes are taking place.
- The basis of Saudi national assessment was primarily formative and summative, but now it has been broadened to include more comprehensible concepts.
- Never before has assessment influenced and driven the curriculum in Saudi Arabia as now.
- Nowadays, the criteria of assessment are a matter of national/governmental discussion. The objectives of teaching and learning have been redefined; the relationship of curriculum to assessment has entered a new partnership, they are no longer the 'master' and 'servant': in other words, the curriculum no longer controls assessment.

Now we have summarised it, we can consider how it is fit for purpose.

3. Fitness for purpose

It is a key concept in assessment that it needs to be fit for purpose. Indeed, assessment is multi-dimensional and if educators fail to realise the various dimensions, they are unlikely to choose the appropriate measurement which best fits that aimed purpose. First of all, any assessment is required to have effective modes and techniques determined ahead of assessment. By 'mode' we mean the overall form of the assessment (general nature, style and character of assessment), whilst the 'technique' means the methods, tools or instruments employed in the assessment; it has a narrower meaning than the mode. A teacher or assessor may use a technique (e.g., multi-choice or essay questions) within different modes. For instance, the teacher may decide to conduct continuous assessment as the mode best suited for one course, while having an end-of-term examination mode for another different course. It is important to point out here that if a teacher is using one or two modes, they may fail to meet some curriculum objectives. In fact, the choice of mode determines different aspects of assessment. For example, it determines when the assessment is going to take place: throughout the course or at the end of term? Who is the assessor: the teacher or an external examiner? What is the general purpose of the assessment; is it mainly to evaluate teaching process, to help the students with their learning or to demonstrate to others that you are doing a good job? In addition, the choice of mode will decide how/what we assess, what we do with the assessment and why we assess. According to Bray (1986), modes can overlap

or merge with each other. The choice of mode will depend on the fitness for purpose which includes the nature of the subject and the objectives of the course.

Moreover, in reality, there should be a coherent assessment policy. Namely, the aims of those who are in charge of conducting the assessment and those who make use of the information in educational institutions should be shared. It is of paramount importance for those educators to shape assessment on agreed guidelines to reach a more robust, supportive and open to innovation system than the traditional one which was based on individual preferences, as the situation was in Saudi Arabia in the 1960s.

In fact, the above discussion has shown that it is of paramount importance for educators in Saudi Arabia as well as in other parts of the world to realise that without having a coherent assessment system, evaluation of the objectives and outcomes can be extremely difficult. Thus, there is much research and investigations still to be done in this area. Rapid developments in assessment may hold the potential to account for these issues, so will be considered in the next section some of the developments in the field of assessment.

3.1 Developments and Future Trends in Assessment

As we have seen above, the last two decades have increased a depth of discussion in the field of assessment, therefore, in addition to a great deal of talk and substantial attempts have been made to introduce new practices. There are different practices which actually reflect different ideologies, but the most salient one is the movement that calls for the recognition of assessment as a way of promoting learning opportunities, rather than a way of catagorising pupils into social roles fit for society.

Furthermore, new technologies and the increasing use of computer-assisted assessment have accelerated the change and opened up the horizon for Saudi as well as others to have new possibilities. The widespread use of computer-adaptive assessment has allowed instructions to be tailored to the appropriate level of the test takers'. Also, it has enabled test takers to be provided with immediate feedback on their performance (Gruba and Corbel, 1997). In fact, researchers are still investigating new ways in which advances in computer technology and linguistic analysis can be incorporated into language testing, including video-mediated testing, automated scoring of open-ended responses, and handwriting and speech recognition, for more details you can refer to: Burstein et al. (1996); J.D. Brown (1997).

Moreover, we should mention here the significant educational initiatives that have led to new trends and purposes in assessment; the most salient is that the curriculum should be relevant to the lives of its receivers (students). It is believed that education should enable pupils to develop their potential to live their lives to the full. Therefore, curriculum and its related assessment need to be broadly utilitarian. In addition, the focus, in terms of learned skills, should not be solely on memory because this does not test fully understanding or encourage critical thinking (Riding, 1990).

In fact, one of the most significant developments in this field is the expansion of recent models of communicative language ability to include other factors such as: aptitude, personality and background knowledge. This would help to recognise the role

of personal characteristics and how they affect performance. Moreover, it would open the way for new development of assessment measures which attempt to build such personal factors into the assessment. Now, as being a teacher, we shall talk about the implication of this aspect for teaching and teachers.

3.2 Implications for Teaching and Learning

Although some educators such as Holt (1969) think that assessment is not useful and does more harm than good, others find it a necessary part of education. We as teachers may say that the main purpose of assessment is to find out what the students have learned, thus we can strategise how to help them to learn more, whilst others claim that assessment operates on greed and fear. Teachers threaten their students into doing what they want done and then handing out the rewards and penalties; consequently, testing arouses fear and satisfies greed (Holt, 1969, p. 52). Fortunately, it is not the case now for most educators and assessors nowadays who seek more flexible ways of measuring students' performance. It is believed that assessment is the means by which learning can be monitored and improved (Bary and Kennedy, 2009). Therefore, it is noted that teachers play a crucial role in assessment; they should be enlightened and prepared for this duty. They should constantly be assessing and adapting adequate strategies and questioning the information obtained through their pupils' results. Teachers should acknowledge both the process and the product of work (Broadfoot, 1991). In addition, they ought to refer to criteria of assessing that are explicit since this could assist students to attain a better understanding and a successful completion of the test. Therefore, it is required for educators to provide their students with a regular and explicit clarification of the assessment criteria (Fair Test, 2007). Furthermore, they should enhance self and collaborative assessment since this gives students the opportunities to produce work that helps to deepen development of their knowledge (Board of Studies NSW, 2006). Teachers are also recommended to provide their students with meaningful feedback and keep records or profiles of achievements handed to them after each semester. In addition, they should spend much time discussing attainment with each pupil and discover ways to help them to be better learners.

Moreover, the assessment is required to take into consideration some important aspects that help students to obtain good performances, for instance:

- Assessment should provide opportunities for students to work together and negotiate required tasks, this would facilitate peer teaching.
- Assessment should also be predominately informal so that the learning process is natural and free of pressure (Brady and Scully, 2005).
- Assessment needs to be sensitive to culture, gender, race, physical disability, and socioeconomic status by applying a range of strategies addressing different human learning methods (NSW Department of Education and Training, 2008).
- Assessment should be formative, continuous and diagnostic. It should provide students and educators with the opportunities to reflect on their practice as well as their overall learning. According to Carol (2002), assessment becomes

formative when the information is used to adapt teaching and learning to meet students' needs.

- Assessment is best when it is ongoing rather than episodic (Victorian Curriculum and Assessment Authority, 2007).
- Lastly but importantly, it is a must for assessment to be fair and reliable because it is necessary to build principle based on validity, otherwise, all the outcomes will be mistaken.

It is of paramount importance to highlight here that it might be difficult to serve three different assessment purposes at the same time. However, it is obligatory for the teacher to recognise the need to balance these purposes knowing which to use and why. Therefore, assessment has become a reflective learning experience for both, students and their teacher, rather than a way of judging.

4. Conclusion

To sum up, it is clear from the above discussion that the continual purpose of the assessment is to provide us with information of others performance, but assessment is not just something done to pupil. It is a way of setting objectives, diagnosing and solving problems, monitoring progress to achieve goals, it is essentially about values and standards. Assessment system in Saudi Arabia as well as in most other countries, is clearly a rapidly changing field which underwent a significant change in the 1990s and will continue doing so. However, it is clearly shown that the former trends in assessment did not appear to fit satisfactory, so there was an urgent need for change. Traditionally, the main purpose for assessment was focusing largely on measuring learning, using the information only to make judegments about students' performance, providing statistics and certification. Nevertheless, the new role of the assessment is to enhance teaching and learning. Currently, assessment tends to be more an evaluation not examination since exams test mainly short-term memory and does not account fully for attainment. Despite the advances in assessment nowadays, there is still a number of influential areas are in urgent need of further investigations, especially since assessment had also inherited from the past some negative issues which needed to be solved, i.e. (results-driven education, or teaching to the test notion). Assessment must respond to the changing needs of society. Furthermore, although research has provided some information of new methods of assessment according to various purposes, the nature and extent of their impact on learning have yet to be fully investigated. I believe that we are still desperately trying to find a convincing answer to this question: Do the syllabuses and assessments provide an adequate framework for improving the quality of teaching and learning? I feel this is a requirement since the ultimate purpose of assessment is to improve students' performance, that an effective assessment should be based on the understanding of how students learn. Indeed, good education should encompass good assessment that serves well defined purposes.

References

1. Ajyaal, (2008). *Teaching English in Saudi Arabia: From Good to Better or from Bad to Worse*. Accessed on 27 December 2014 from http://ajyaal.ma3ali.net/articles-action-show-id-535.htm.

2. Board of Studies NSW (2006). General Principles for Planning, Programming, Assessing, Reporting and Evaluating. Accessed on 29 December 2014 from http://arc.boardofstudies.nsw.edu.au/go/k-6/assess-principles.

3. Bond, L. (1995). Unintended consequences of performance assessment: Issues of bias and fairness. *Educational Measurement: Issues and Practice*, 14(4): 21-24.

4. Borg, S., & Edmett, A. (2018). Developing a self-assessment tool for English language teachers. *Language Teaching Research*, 10, (1177): 1-25.

5. Boston, C. (2002). The concept of formative assessment. *Practical Assessment, Research & Evaluation*, 8(9): 624-637.

6. Brady, L. & Kennedy, K. (2009). Celebrating Student Achievement: Assessment and Reporting (3rd Ed.). Sydney: Pearson.

7. Brady, L. & Scully, A. (2005). Engagement inclusive classroom management. Sydney: Pearson.

8. Broadfoot, P. (1991). *Assessment: A Celebration of Learning*. Canberra: ACSA.

9. Carol, B. (2002). The concept of formative assessment. *Practical Assessment Research & Evaluation*, 8(9).

10. Corbett, H.D., & Wilson, B.L. (1991). *Testing, Reform and Rebellion*. NJ: Norwood.

11. Damico, J. (1992). "Performance Assessment of Language Minority Students." In *Focus on Evaluation*, volume 1.

12. Department Of Education (2008). Learning, Teaching and Assessment Principles. Accessed on 2 January 2015 from http://www.education.tas.gov.au/curriculum/ltap.

13. Eisner, E. (2001). The Educational Imagination: On the Design and Evaluation of School Programmes (3rd Ed.). Prentice Hall: Macmillan.

14. Fair Test (2007). Principles and Indicators for Student Assessment Systems. Accessed on 17 January 2015 from http://www.fairtest.org/principles-and-indicators-student-assessment-syste.

15. Groundwater-Smith, S., Ewing, R. & Le Cornu, R. (2007). *Challenges and Dilemmas* (3rdEd.). Sydney: Thomson.

16. Holt, J. (1969). *The Underachieving School*. London: Pitman.

17. Horton, T. (1990). *Assessment Debates.* New York: The Open University Press.

18. Jones, R. and Bray, E. (1986). *Assessment from Principles to Action*. London: Macmillan.

19. Masters, G. & Forster, M. (2000). *The Assessments We Need*. Melbourne: Australian Council for Educational Research.

20. McNamara, T. (2001). Language assessment as social practice: Challenges for Research. *Language Testing*, 18: 333-349.

21. Newstead, S. E. (2004). Time to make our mark. *The Psychologist, 17: 20-23*

22. NSW Department of Education and Training (2007). Principles for Assessment and Reporting in NSW Government Schools. Accessed on 3 January 2014 from http://www.schools.nsw.edu.au/learning/k-6assessments/principles.php.

23. O'Sullivan, B. 2006. Testing Language for Business: A Critical Overview of Current Practice. Accessed on 14 January 2014 from http://www.hltmag.co.uk/jul07/sart04.htm ESP Malaysia, 3: 17-31.

24. Riding, R.J. and Butterfield, S. (1990). *Assessment and Examinations in the Secondary School*. London: Guildford and King's Lynn.

25. Victorian Curriculum and Authority (2007). Assessment Principles. Accessed on 5 January 2017 from http://vels.vcaa.vic.edu.au/assessment/assessmentprinciples.html.