Enhancing Malaysian Teachers' Assessment Literacy

Lim Hooi Lian[1], Wun Thiam Yew[1] & Chew Cheng Meng[1]

[1] School of Educational Studies, Universiti Sains Malaysia, Penang, Malaysia

Correspondence: Lim Hooi Lian, School of Educational Studies, Universiti Sains Malaysia, 11800 Penang, Malaysia. E-mail: hllim@usm.my

Abstract

Currently, in order to reform the Malaysian education system, there have been a number of education policy initiatives launched by the Malaysian Ministry of Education (MOE). All these initiatives have encouraged and inculcated teaching and learning for creativity, critical, innovative and higher-order thinking skills rather than conceptual knowledge, procedural knowledge, and rote memorization. The reform in teaching and learning approaches in Malaysian teacher education should also be reflected in the method of assessment as assessment is seen as a vital part of instruction in the culture of learning. In view of the need for changing school assessment culture, teachers' assessment literacy becomes one of the main concerns. Assessment literacy is regarded as the sound knowledge and skills in educational assessment required by teachers in assessing students' learning outcome. This article presents practical suggestions concerning knowledge and skill target of teachers' assessment literacy. Assessment literacy should not be fully addressed in theoretical fashion, but in a more practical and real-life manner. The five knowledge and skill targets that teachers should seriously focus on are: 1) validity of assessment, 2) reliability of assessment, 3) transparency of assessment, 4) fairness of assessment, and 5) using of assessment information.

Keywords: assessment literacy, validity, reliability, fairness, transparency

1. Introduction

Currently, in order to reform the Malaysian education system, there have been a number of education policy initiatives launched by the Malaysian Ministry of Education (MOE) such as Primary School Standard Curriculum (*Kurikulum Standard Sekolah Rendah* KSSR, MOE, 2012a) and Education Development Master Plan (*Pelan Induk Pembangunan Pendidikan*, MOE, 2012b). All these initiatives have encouraged and inculcated teaching and learning for creativity, critical, innovative and higher-order thinking skills rather than conceptual knowledge, procedural knowledge and rote memorization. Thus, Malaysian teachers had been motivated to move towards student-centered learning and constructivist learning approaches in order to develop students' skills in creative, innovative, critical, competent, competitive, progressive, learning to know, learning to do, learning together and learning to be as desired by the nation (MOE, 2012).

However, the reform in teaching and learning approaches should be reflected in the method of assessment as assessment is seen as a vital part of instruction in the culture of learning (Shepard, 2000). Methods of assessment must be developed that better reflect current understanding about the ways in which knowledge and skill are constructed through learning. In short, "as the curriculum changes so must the assessment" (National Council of Teachers of Mathematics, 1989). Research results had revealed that traditional assessment methods especially paper and pencil test should be replaced, or at least complemented by alternative assessments such as performance-based assessment, outcome-based assessment and student-centered assessment (Shepard, 2000). In view of the need for changing the school assessment culture, the Malaysian Ministry of Education has provided school teachers with various supports such as training programs (i.e. workshops, briefing, forums, conferences, etc) and references sources (i.e. module, guide book etc). However, one of the main concerns is still in the area of teachers' assessment literacy. Assessment literacy is regarded as the sound knowledge and skills in educational assessment required by teachers in assessing students' learning outcome. Stiggins (1999) stated that assessment literates are able to identify the differences between sound and unsound assessment.

Popham (2011) stated that there are two prominent reasons why school teachers need to become more knowledgeable regarding educational assessment. The first reason is that today's educators are being assessed

primarily on the basis of students' scores on accountability test; hence it is 'flat-out folly' for new teachers not to understand the tools being used to judge them. The second reason is that educational assessment had been proven play an important role in substantially enhancing and promoting the quality of students' learning if properly created and skillfully employed.

Meanwhile Webb (2002) voiced the opinion that there are two causes why the issue of teachers' assessment literacy has received increasing attention over the past decade. First, there is significant important and greater acceptance of using various methods of assessment for the purpose of formative and summative assessments. Second, the emergence of a student-centered learning reform has made expectations for students to learn more explicit. Both causes have great impact on the more formal use of alternative assessment to identify students' learning expectation. They have increased the need and expectation on teachers to understand how student learning can be adequately assessed, to determine whether students have achieved the learning expectations and what meaning should be given to the assessment information gathered. Directly, it highlights the important of alignment of concepts in Constructive Alignment Theory (Biggs, 2003). That is, teachers play a vital role in deciding assessment methods that will tell them appropriately and precisely, how well individual students have attained the intended learning outcomes. Also, this information is pivotal for teachers to help and improve their students' quality of learning.

Therefore, in this article, I would like to lay out some knowledge and skill targets concerning to teachers' assessment literacy that could be discussed in more practical and real-life ways rather than just in theoretical manner, in the process of implementing educational assessment particularly classroom assessment.

2. Validity of Assessment

Validity and reliability play the important role in the success of any assessment. Malaysian Examinations Syndicate has set up a mechanism of quality assurance to ensure that the implementation of school-based assessments is in accordance to its intended goals. The proposed quality assurance mechanism includes the mentoring, monitoring, moderating, and detecting measures. Teachers will be guided, facilitated, and mentored in the process of conducting, recording, and reporting of an assessment (Malaysian Examinations Syndicate, 2014a). Popham (2006) has so expressively mentioned, "*Validity and reliability are the meat and potatoes of the measurement game.* Both are essential principles in the educational measurement area. Therefore, every school teacher should know, understand and put into practice such conceptual essentials in order to make a better assessment decision on students learning and teaching (MOE, 2009).

Validity is considered an evaluative judgment about the degree to which the assessment results are appropriate for making certain educational inference and decisions (Messick, 1993). To achieve it, teachers need to be capable of deciding assessment methods and constructing assessment tasks that appropriately assess the intended learning outcomes to be assessed. But still too often, Malaysian school teachers do not succeed in this intention. They assessed what they can assess and what had been covered in the class (Mohamad, 2006; Salbiah, 1995). In other words, they emphasize more on assessing the content of syllabus rather than students' learning outcomes. If the teachers assess what the students are able to know and to do with what they process from what they think, it would more meaningful. In other words, teachers need to assess students' performance in the contexts which are as close as possible with the intentions lying behind the learning outcomes in the first place.

Over the years, there has been strong agreement amongst some writers (Gronlund, 2006; Lihanna, 2003; McMillan, 2008) that it is important to vary assessment methods so that the same students are not discriminated against repeatedly because they are not proficient in certain methods of assessment that had been over-used by teacher, especially paper and pencil tests. Furthermore, the affective and psychomotor domain of learning normally should be assessed by using performance-based assessment methods, not just where students are writing about such skills and knowledge in examination halls. For instance, when attempting to assess manipulative skills, the assessment should not be fully dependent on the quality and ability of producing written responses, but also on the quality of the performance such as hands-on activities and demonstrations.

After deciding the assessment method to be used, teachers need to follow certain fundamental principles and standard guidelines suggested by assessment experts in creating more quality and effective assessment tasks. However, this process has always been ignored and sidelined by school teachers. As a result, many assessment tasks are poorly designed and ineffectively used (Suah, 2012). Researchers have revealed that many teachers are ill-prepared to develop various methods of assessment and to create assessment tasks, especially authentic assessment (e.g., Bol, Stephenson, O'Connell, & Nunnery, 1998; Stiggins & Conklin, 1992; Wiggins, 1989). Teachers who are less skilled and less prepared in constructing assessment tasks, perceive these to be more challenging and difficult than constructing traditional paper-and-pencil tests. Many teachers were found not to be

good judges of their own assessment tasks in terms of the quality and effectiveness (Bol & Strage, 1996).

Therefore, teachers need to become more knowledgeable regarding the selection of assessment methods and development of assessment tasks so they can arrive at justifiable inferences about students' covert skills and knowledge. The justifiable inferences and evidences then play the prominent role in the making of arguments to support the accuracy of interpretations and conclusion about students' achievement. As Popham (2009) states teachers who are genuinely assessment literate will not only efficiently develop more appropriate assessments, but also will become familiar with the various potential assessment methods. The more accurate the assessment information that teachers gather, the more appropriate of the interpretation and inference of the assessment results that bring to the better degree of validity. In short, school teachers need to possess adequate assessment literacy so the assessment information they gather is at least substantially, appropriately or significantly better. Quitter (1999) revealed that teachers should understand the relationship between validity and the principles of educational assessment, in as much as it is an exemplar of the appropriate collect and use of assessment information. Moreover, collections of different types of validity evidence are fundamentally different. However, at least, the classroom teachers have to understand that the chief kind of validity evidence they need to attend to should be content-related.

3. Reliability of Assessment

The reliability of an assessment refers to the extent it consistently and accurately assesses students' learning. When the results are reliable, it can be concluded that repeated or equivalent assessments will provide the consistent results. This will put teachers in a better position to make generalized statements about students' levels of performance, which is especially important when teachers use the assessment results to make decisions about classroom teaching and learning.

Our new assessment culture especially the implementation of school-based assessment, aims to develop learners' potential comprehensively and holistically. It is hoped to achieve the aspiration of National Philosophy of Education towards developing learners' physical, emotional, spiritual, and intellectual abilities (Malaysian Examinations Syndicate, 2014b). This has led to a strong interest in various method of performance-based assessment (such as presentation, demonstration, simulation and discussion) which involve observation judgments. However, the issue of whether the observation of complex behavior can be carried out in a convincing and trustworthy manner occurs. Linn, Baker, and Dunbar (1991) and Moss (1992) state that performance assessment generally faces the problem of lower value reliability by teachers because it is not high-standardized procedures. Mehrens (1992) noted that there are several threats to the value reliability in performance assessment. First, it has to be implemented with independent observations. Second, it has to do with the subjective nature of the scoring process.

Besides, Popham (2009) revealed that although certain formats of performance assessment tasks (such as essay, project writing, report and problem solving task) always provide particularly informative evidence about students' skills and knowledge, but the scoring of students' responses can often be erratic due to the problem of quality of the judgmental procedures.

Therefore, scoring of students' responses to the various formats of assessment tasks should be based on a well-formed scoring rubrics or checklist. Teachers need to have knowledge and skill on developing and using scoring rubrics or checklist professionally. The assessment criteria and scoring rubrics should have good inter-rater reliability (when more than one examiner marks the students' work), as well as good intra-rater reliability (examiner should come up with the same results when marking the same work on different occasions); so students' performances can be accurately appraised.

4. Transparency of Assessment

By implementing school-based assessment, Malaysian assessment attention has shifted to the implementation of formative assessment approaches. It is claimed that the implementation should involve in helping students to possess the concept of the standard (such as the learning outcome and criteria assessment) being aimed for. This standard is important for students to be able to judge the quality of what they are producing and be able to regulate what they are doing during the process (Sadler, 1989). However, this practice is always side-lined in Malaysian assessment system. The final grade of a students' achievement is still the main goal. Most of the teachers prefer producing and assuring students to get a good grade as they believe that the final product is still the main determiner to the success in teaching and learning. James, Black, McCormick, Pedder, and William (2006) state that although this main characteristic of formative assessment is already widespread in the secondary education, but is inferred very superficially, with an irresistible focus on award achievement, far from encouraging an orientation towards student autonomy and 'learning how to learn'. Rashidah (2004) revealed that

most of Malaysian school teachers faced the problem and confusion to understand the purpose and concepts of formative assessment. They regarded formative assessment as an assessment practice that seeks to prepare students for public examinations.

As a consequence, there is an increasing acceptance of the need for a greater transparency in assessment. Move should be made to ensure the assessment process and assessment criteria are very clear to all students. 'Transparency' refers to the extent the students' understand where the goalposts are (Race, 2009). The goalposts in this context refer to the match between intended learning outcomes and assessment criteria. The assessment criteria describe the standards to which the intended learning outcomes are to be performed by students, and also be seen as an indication of achievement; specifying the forms in which students will present evidence of their achievement of the outcomes.

Race (2009) asserts that the intended learning outcome to be assessed and assessment criteria should have no hidden agendas. Students should not be playing the game 'guess what's in our teachers' minds' or 'guess what will be tested by our teachers'. Assessment should relate closely to the intended learning outcomes as mentioned in text books and syllabus documentation. Besides, the links between intended learning outcomes and the assessment criteria should be clearly seen (not just by administration officer or head department, but by students themselves). As Sadler (2005) states that the judgment of the quality of student's will be made based on assessment criteria. Thus, students merit knowing and understanding it. A primarily prospective purpose of assessment criteria is to enable students to use the information to shape and improve their work intelligently and appropriately while it is being developed.

5. Fairness of Assessment

Fairness of assessment means the assessment should allow for all students of different backgrounds such as socioeconomic status, genders and ethnics to do equally well. All students should have an equal opportunity to perform their skills and knowledge being assessed. In other words, all students should have equivalence of opportunities to succeed even if their experiences are not identical (Race, 2009).

It is especially crucial that all assessment tasks should be seen to be fair by all students. For an assessment task to be fair, teacher should ensure that its content, format, context, and performance expectations reflect students' knowledge, skills, values and experiences that are equally familiar and appropriate to all of them. In other words, it must be as free as possible of socioeconomic status, cultural, ethnic, and gender stereotypes. Bias in a task is similar to the idea of extraneous interference. For instance, if an assessment task is created in the context of hockey and students who have a knowledge or skill of hockey have an advantage on the task, that knowledge and skill is an extraneous factor. Bias, however, refers to things that systematically affect entire groups of students rather than individual students. The context of task becomes a biasing factor if particular groups of students (e.g. female students) know less about hockey than other groups of students (e.g. male students). As Linn, Baker, and Dunbar (1991) revealed that gaps in performance among groups exist because of difference in familiarity, exposure, and motivation on the tasks of interest.

The format of assessment task also can differentially affect results for different groups. For instance, females are likely to outperform than males on the open-ended item, predominantly when this form of task involving personal response. However, the gap between genders becomes narrows if the task format is selected-response test item or short answer item (Gipps & Murphy, 1994). Due to this, one of the main purposes of diversifying assessment method may be encouraged; that is, the use of a range of assessment methods and modes so that those who are disadvantaged on certain assessment methods have an opportunity to offer alternative evidence of their performance (Linn, 1992).

Besides, issue of fairness occurs not only in the selection of assessment method but also in the scoring of responses (Sackett, 1987). As Stiggins (1987) has stated, it is critical that the scoring procedures are designed to assure that performance ratings reflect the examinee's true capabilities and are not a function of the perceptions and biases of the persons evaluating the performance. The professional training and calibrating of raters are vital and critical in this regard.

6. Using of Assessment Information

Assessment information can be used in many ways. Teachers can use it to help students in learning, track students' progress towards important learning outcome, provide information to parents and administration department, or do the class streaming. In general, all the assessment information uses can be categorized into two, assessment of learning and assessment for learning (Stiggins, Arter, Chappuis, & Chappuis, 2007). Both categories have their place in classroom and education. Malaysia has been doing both in the educational

assessment system. However, the teachers' understanding of both, especially assessment for learning is still too narrow (Rashidah, 2004). Thus, we need to expand their understanding of its role that should play to maximize student achievement while minimize side effects for students.

A prime focus of assessment for learning is to identify areas that may need improvement, do the judgments about the quality of student responses (performances or works) that can be used to shape and improve the student's competence and skill (Sadler, 1989). Assessment of learning contrasts with assessment for learning as it takes place after the learning has been completed and provides information and feedback that sum up the teaching and learning process. It is concerned with summarizing the strengths, weaknesses of a student. It describes the extent to which a properly implemented subject has attained its goals or objectives. It is geared towards reporting at the end of a subject learning especially for purposes of certification. Giving report cards and grade are the most common examples of reporting the student's achievement in assessment of learning. It is basically passive and does not normally have immediate effect on learning, although it often impacts decisions which may have profound educational and personal consequences for the student (Sadler, 1989).

Assessment for learning typically is not graded and acts as a gauge to students' learning progress and to identify instruction effectiveness. This type of assessment has allowed teachers to "rethink" and then "re-deliver" the teaching approach and content, to confirm their students are on the right track. Besides, it is a good practice to integrate this type of assessment to "assess" students' knowledge and skill before expecting them to perform well in assessment of learning, such as semester examination.

However, school teachers often worry that they don't have enough time to assess students along the way. They worry about not having enough time to prepare the assessment tasks, conduct the assessment and mark the students' responses. They rush to cover the syllabus within the time frame (Aidarwati & Abdul, 2013).

Actually, without time to do the reflection on students' learning ability, interact meaningfully with assessment information to prepare more effective lesson, students are unlikely to retain much of what is "covered" in the classrooms. Students are actually learning less as the teachers are teaching the syllabus, not the students.

In addition, assessment for learning does not have to take an inordinate amount of time. There are many methods of assessment that are simple and easy to implement on a daily basis, for instance observation, quiz, presentation, demonstration (manipulative skill), oral test (language subject). On balance, the time they take from one or two lessons are well worth the information that the teachers gather and the retention students' gain. If the assessment method involves individually, for instance, oral test or interview which is hard to conduct it concurrently, teacher may conduct it informally; do the random selection of students involved for different sessions of assessment.

Guskey (2007/2008) suggested that assessment for learning to be the essential part of the teaching process. Teachers should change their approach in three important ways. They must: 1) use assessments information intelligently and professionally for the benefit of teachers and students, 2) follow assessments fundamental principles with high-quality corrective instruction, and 3) give students second opportunities to show their success.

By implementing assessment for learning, teachers must take action based on the assessment result. They will be able to help their students to achieve success in learning by improving their teaching approach based on the assessment result that they have assembled. The teachers need to ask themselves, "Which students should pay more attention now? Which students should provide different teaching approaches to understand better of my lesson? Which students are not learning anything new, because I haven't challenged them?" Teachers must be ready to prepare both remediation and enrichment activities for those who need them. In the simple words, the follow-up (such as enrichment and remediation activities) and corrective teaching approach must present in a new way and involve students in a more meaningful and interesting learning experiences that are more appropriate for them (Guskey, 2007/2008).

7. Conclusion

Apparently, assessment literacy is a commodity needed by teachers for their own long-term well-being, and for the educational well-being of their students. However, previous studies revealed that classroom teachers are unprepared to sufficiently assess their students (e.g., Mertler, 1999; Plake, 1993). In Suah's study which involved 3866 teachers from primary and secondary schools, revealed that majority of Malaysian teachers' literacy assessment level was less satisfactory. Most of them stated that their last time of attending the assessment training program was five or more than five years ago. The teachers believed that they had not received adequate formal training in their undergraduate education programs and school professional development program (Suah, 2012). In this regard, four suggestions that had been discussed critically by Koh and Velayutham (2009) towards

improving teachers' assessment literacy have been found to be a good reference in developing and advancing educational assessment in our country.

(a) The effort to actualize assessment literacy reform among the school teachers will not happen in a single or sporadic workshop. It is a long-term process to be followed. Hence, teachers' assessment literacy can only be enhanced through high-quality, comprehensive, and ongoing professional development.

(b) Professional development of assessment literacy should be part of the daily practice of teachers. It should no longer be seen as an ad hoc event that happens only on a few days of the workshop or briefing. Therefore, before changing the old practice of assessment culture in our school, teachers should make a first move. They should change their old beliefs and update their assessment practices from time-to-time. They need to collaborate actively through school professional development communities. This will change assessment culture in the long term.

(c) Teachers need to be encouraged and guided to take the lead in redesigning assessment process based on the fundamental principle of assessment. It is vital to produce more quality, reliable and valid assessment result. In the long run, the reform of their assessment practices will be beneficial to both teacher development and student learning.

(d) In addition to design more quality assessment tasks to assess student's knowledge and various thinking skill, teachers also should be capable in using the assessment result to improve a student's learning, for instance giving more effective and timely feedback which is tailored to the student's strengths, weaknesses and understandings. It is the best way in telling a student about what s/he is doing well and what needs to be improved.

Acknowledgements

References

Aidarwati, M. B., & Abdul, G. A. (2013). A comparative analysis of primary and secondary school teachers in the implementation of school-based assessment. *Malaysian Journal of Research*, *1*(1), 28-36.

Biggs, J. B. (2003). *Teaching for quality learning at university*. Buckingham: Open University Press/Society for Research into Higher Education.

Bol, L., & Strage, A. (1996). The contradiction between teachers' instructional goals and their assessment practices in high school biology courses. *Science Education*, *80*, 145-163.

Bol, L., Stephenson, P. L., O'Connell, A. A., & Nunnery, J. A. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *Journal of Educational Research*, *91*, 323-330.

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

Gronlund, N. E. (2006). *Assessment of student achievement* (8th ed.). Boston: Pearson Education, Inc.

Guskey, T. R. (2007/2008). The rest of the story. *Educational Leadership*, *64*(4), 28-35.

James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, *17*(2), 109-138. http://dx.doi.org/10.1080/09585170600792712

James, M., Black, P., McCormick, R., Pedder, D., & Wiliam, D. (2006). Learning how to learn, in classrooms, schools and networks: Aims, design and analysis. *Research Papers in Education*, *21*(2), 101-118. http://dx.doi.org/10.1080/02671520600615547

Kementerian Pelajaran Malaysia [Ministry of Education]. (2009). *Standard Guru Malaysia* [Malaysian Standard Teachers]. Putrajaya: Kementerian Pelajaran Malaysia.

Kementerian Pelajaran Malaysia [Ministry of Education]. (2012a). *Primary School Standard Curriculum [Kurikulum Standard Sekolah Rendah]*. Retrieved May 15, 2014, from http://www.moe.gov.my/v/soalan-lazim-view?id=146&cat=30&keyword=&page=1&

Kementerian Pelajaran Malaysia [Ministry of Education]. (2012b). *Pelan Pembangunan Pendidikan Malaysia 2013–2025 [Education Development Master Plan]*. Retrieved May 15, 2014, from http://www.moe.gov.my/v/pelan-pembangunan- pendidikan-malaysia-2013-2025

Koh, K. H., & Velayutham, R. L. (2009). Improving teachers' assessment literacy in Singapore schools: An analysis of teachers' assessment tasks and student work. Retrieved May, 13, 2014, from ww.nie.edu.sg/

files/NIE_research_brief_09_002.pdf

Lembaga Peperiksaan Malaysia [Malaysian Examinations Syndicate] (2014a). *Panduan penjaminan kualiti* [*Quality assurance guidelines*]. Retrieved on July 17, 2014, from http://apps2.moe.gov.my/lponline/v1/files/penerbitan/pbs/PANDUAN%20PENJAMINAN%20KUALITI%202014_18062014%20.pdf

Lembaga Peperiksaan Malaysia [Malaysian Examinations Syndicate]. (2014b). *Panduan pengurusan PBS (Pentaksiran Berasaskan sekolah)* [*Management of school based assessment guidelines*]. Retrieved July 17, 2014, from http://apps2.moe.gov.my/lponline/v1/files/penerbitan/2014/Panduan_Pengurusan_PBS_21_April_2014 2.pdf

Lihanna. (2003). Penilaian autentik dalam pendidikan awal kanak-kanak melalui portfolio [Authentic evaluation in early childhood education through portfolio]. *Masalah Pendidikan*, *26*, 13-20.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex performance-based assessment: Expectations and validation criteria. *Educational Researcher*, *20*(8), 15-21.

Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education*, *21*(2), 133-149. http://dx.doi.org/10.1080/02671520600615638

McMillan, J. H. (2008). *Assessment essentials for standard-based education* (2nd ed.). California: Corwin Press.

Mehrens, W. A. (1992). Using performance assessment for accountability purposes. *Educational Measurement*: *Issues and Practice*, *2*(1), 3-9.

Mertler, C. A. (1999). Assessing student performance: A descriptive study of the classroom assessment practices of Ohio teachers. *Education*, *120*(2), 285-296.

Messick, S. (1993). Validity. In R. L. Linn (Ed.), *Educational measurement* (2nd ed., pp. 13-104). Phoenix: American Council on Education and Oryx Press.

Mohamad, A. M. A. (2006). *Amalan pentaksiran di sekolah menengah [Assessment pratice in secondary schools]* (Unpublished doctoral thesis). Universiti Malaya, Petaling Jaya.

Moss, P. A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. *Review of Educational Research*, *62*(3), 229-258.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematic*s. Reston, VA: Author.

Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher*, *6*(1), 21-27.

Popham, W. J. (2006). *Assessment for Educational Leaders*. Boston, MA: Pearson.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, *48*, 4-11. http://dx.doi.org/10.1080/00405840802577536

Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator*, *46*, 265-273. http://dx.doi.org/10.1080/08878730.2011.605048

Quilter, S. M. (1999). Assessment literacy for teachers: Making a case for the study of test validity. *The Teacher Educator*, *34*(4). 235-43.

Race, P. (2009). *Designing assessment to improve Physical Sciences learning. A Physical Sciences Practice Guide*. Retrieved May 13, 2014, from http://www.heacademy.ac.uk/physsci

Rashidah, B. H. (2004). *Kesediaan guru melaksanakan pentaksiran untuk pembelajaran [Teachers' readiness in implementing assessment for learning]*. (Unpublished master education thesis), University of Malaya, Petaling Jaya.

Sackett, P. R. (1987). Assessment centers and content validity: Some neglected issues. *Personnel Psychology*, *40*, 13-25.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems, *Instructional Science*, *18*, 19-144.

Sadler, D. R. (2005) Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, *30*(2), 175-194.

Salbiah, I. (1995). *Amalan penilaian formatif dalam pelaksanaan Penilaian Kemajuan Berasaskan Sekolah oleh*

guru Bahasa Melayu dan Matematik [*Formative assessment practices in the implementation of the School-Based Evaluation of the Malay Language and Mathematics teachers*] (Unpublished master education thesis). UniversitiTeknologi Malaysia, Skudai.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, *29*(7), 4-14.

Stiggins, R. J. (1987). Design and development of performance assessments. *Educational Measurement: Issues and Practice*, *6*(3), 33-42.

Stiggins, R. J. (1999). Are you assessment literate? *The High School Journal*, *6*(5), 20-23.

Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment.* Albany: State University of New York Press.

Stiggins, R. J., Arter, J. A., Chappuis, J., & Chappuis, S. (2007). *Classroom assessment for student learning: Doing it right–Using it well*. New Jersey: Pearson Education, Inc.

Suah, S. L. (2012). *Analisis model literasi dan amalan pentaksiran guru sekolah serta kajian tentang jurang antara keduanya* [*Analysis of assessment literacy and practice school teacher and study of the gap between them*] (Unpublished doctoral thesis). Universiti Sains Malaysia, Penang.

Webb, N. L. (2002). *Assessment literacy in a standards-based urban education setting*. Retrieved May 13, 2014, from http://www.wcer.wisc.edu/archive/mps/AERA2002/Assessment%20literacy%20NLW%20Final%2032602.pdf

Wiggins, G. (1989, May). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan*, *70*, 703-713.

Copyrights