

TUNKU AHMAD, ZUBAIRI, IBRAHIM, OTHMAN, RAHMAN, RAHMAN, NORDIN & NOR:
ASSESSMENT FOR LEARNING PRACTICES AND COMPETENCY AMONG MALAYSIAN UNIVERSITY
LECTURERS: A NATIONAL STUDY

**Assessment for learning practices and
competency among Malaysian university
lecturers: a national study**

Practitioner Research
In Higher Education
Copyright © 2014
University of Cumbria
Vol 8 (1) pages 14-31

Tunku Badariah Tunku Ahmad*, AinolMadziahZubairi*, MohdBurhan Ibrahim*, Joharry Othman*, NikSuryaniAbd Rahman*, ZainurinAbd Rahman*, Mohamad SahariNordin*, ZainabMohd Nor**

*Institute of Education, International Islamic University Malaysia (IIUM), Kuala Lumpur

**Faculty of Education, UniversitiTeknologi Malaysia (UiTM), Shah Alam

tbadariah@iium.edu.my

Abstract

This paper reports the findings of a national study involving 534 lecturers from 33 higher learning institutions in Malaysia to find out their self-reported practices and perceived competencies in assessment for learning. Data were collected using a 24-item assessment practice inventory drawn from five of the six standards stipulated in an established US-developed assessment competency framework for teachers. Principal Component Analysis (PCA) and descriptive statistics were used to analyse lecturers' reported practices and competencies. The PCA procedures extracted four underlying dimensions of AfL practice among the lecturers that corresponded to four of the five standards used. Further descriptive analysis produced a rank order of the four AfL dimensions in terms of practice and competency. Lecturers reported to quite frequently practise 'communicating results and feedback to students' and 'using diverse AfL methods'. They also perceived to be more competent in these two aspects than in AfL-compliant grading practices and recognising unethical uses of assessment information. The findings produced adequate empirical support for instructional interventions and training to be provided at the national level to upgrade lecturers' current competencies in assessment and support their adoption of AfL.

Keywords

Assessment for learning; assessment practices; assessment competency; assessment standards; assessment practice inventory; Malaysian higher education.

Introduction

Assessment is an integral component in any instructional context and educational setting. In higher education, it is pervasive and plays at least two key roles; one in ensuring institutional quality and accountability, and the other, in improving student learning (Ewell, 2009). University lecturers' use of assessment affects the depth and quality of what students learn, their choice of learning strategies, how they manage their study time (Australian National Training Authority, 2002; Brown, Bull and Pendlebury, 1997), and quite significantly, their motivation to continue learning (Boud, 1995; Harlen and Crick, 2003). Thus, the importance of assessment cannot be

Citation

Tunku Ahmad, T.B., Zubairi, A.M., Ibrahim, M.B., Othman, J., Rahman, N.S.A., Rahman, Z.A., Nordin, M.S. & Nor, Z.M. (2014) Assessment for learning practices and competency among Malaysian university lecturers: a national study, *Practitioner Research in Higher Education Journal*, 8(1), January, 14-31.

overstated. In fact, according to Anderson (2004) and Hannafin *et al.* (2003), no factor influences a learning environment as much as assessment.

Assessment has been defined variously by different scholars and authorities, for example as 'any processes that appraise an individual's knowledge, understanding, abilities or skills' (The Quality Assurance Agency for Higher Education, 2011:3), 'vehicles for gathering information about students' achievement or behavior' (Marzano, 2000:86), and 'an ongoing process aimed at understanding and improving student learning' (Angelo, 1995:7). The first two definitions constitute what theorists say are assessment *of* learning (AOL), which in essence is a certification of what and how much students have acquired over the course of learning. The third definition contains an element of assessment *for* learning (AfL) where the central purpose of assessment, which is to stimulate greater student progress and achievement, is included. Stiggins (2002) makes a distinction between these two assessment types and remarks that both are essential and have a place in education. However, in order to maximize student achievement, he suggests paying far greater attention to AfL.

Black and Wiliam (1998) define AfL as a formative assessment system in which continuous and informative feedback is used to modify teaching to meet students' learning needs and galvanize them to learn more. Stiggins (2002) agrees with this notion and refers to AfL as the use of classroom assessment process and the continuous flow of information about student achievement to further advance their learning. Another key theorist, Klenowski (2009:264), offers the following definition:

Assessment for Learning is part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance ongoing learning.

In agreement with Klenowski (2009), Black *et al.* (2004:10) summarize AfL as:

Any assessment for which the first priority in its design and practice is to serve the purpose of promoting students' learning. An assessment activity can help learning if it provides information that teachers and their students can use as feedback in assessing themselves and one another and in modifying the teaching and learning activities in which they are engaged. Such assessment becomes 'formative assessment' when the evidence is actually used to adapt the teaching work to meet learning needs.

Central to AfL are the ideas of formative assessment, giving feedback to students in ways that enable them to improve, and providing feedforward to help them plan their next learning steps. It is quite clear that the fundamental purpose of AfL is not to merely check on students, but rather to help them learn more and acquire greater potential. Therefore, AfL is a catalyst for reform in instructional practices as it develops students' self-regulated and autonomous learning ability (Sahari, 1999). These two are critical skills for lifelong learning in the 21st century's knowledge based economy (Song and Koh, 2010). AfL also bridges between theory and practice (Riley and Stern, 1998), and creates 'a shared academic culture dedicated to assuring and improving the quality of higher education' (Ellyn, 2000:2).

Assessment in Malaysian Higher Education: The National Context

In Malaysia, calls for the accountability of higher education in promoting student learning are a national agenda that is slowly gaining momentum. In this regard, Malaysian universities are required to give paramount importance to two key components in the structure of academic programmes. They are: (i) clear and measurable programme learning outcomes, and (ii) quality assessment that is well-aligned with the intended outcomes. The relationship between these two components is an intricate one where the provision of clear and measurable learning outcomes is expected to guide the use and practice of assessment among university lecturers. The requirement to incorporate measurable outcomes and well-aligned quality assessment into higher education academic programmes is mandated by two sources: (i) the Malaysian Ministry of Higher Education through its Quality Assurance Division, and (ii) the National Higher Education Strategic Plan (2007-2020). Since its articulation, the mandate has stimulated some public universities to formulate a policy to guide their teaching staff's assessment practices, and to embark on a series of intensive student assessment training for the staff.

However, it has not had the same effect on many other public and private universities in the country. To date, there are still many that do not have a clear policy on assessment or provide necessary training for lecturers, hence leaving them clueless as to the 'what' and the 'how' of student assessment. Some insight into this situation was given in Zubairi, Sarudin and Nordin (2008), who reported that Malaysian university lecturers' use of assessment tends to be restricted to paper-and-pencil tests with quizzes and traditional formats (such as multiple-choice, true-false and essay questions) being the most widely and frequently utilised methods. The use of alternative and authentic assessments (e.g. observations, demonstrations, portfolios, etc.) was found to be uncommon among the lecturers. The study was conducted at a single public university involving a sample of 135 lecturers, and the results are hardly generalisable to other higher education institutions in Malaysia. However, the adherence to traditional formats reported in the study is known to be a long standing practice, and provides some indication as to what Malaysian university lecturers commonly understand to be student assessment. It also suggests that their assessment knowledge and practice have not improved much over the years. Ong (n.d.) noted that the adherence to paper-and-pencil tests is also prevalent among Malaysian school teachers, a practice they most likely adopted from years of schooling and being trained in an education system that is deeply entrenched in traditionalist pedagogical and assessment methods.

A recent 2012-2013 initiative by the Malaysian Higher Education Leadership Academy (locally known by its Malay acronym, AKEPT), a division within the Ministry of Higher Education, attempts to address the lack of assessment knowledge and competency among university lecturers. Faculty's predominant, almost sole reliance on traditional assessment methods and assessment of learning practices was well-noted and factored in. Hence in collaboration with a UK-based expert, the academy embarked on a project to develop a set of modules in learning assessment, which it completed and successfully produced in early 2013. The modules were then used in subsequent training and workshops on student learning assessment for university representatives with the expectation that the expertise gained will be cascaded to other lecturers of the same university.

AKEPT's modules start off with the notion of assessment for learning (AFL) and multiple

strategies to implement it. This approach was adopted in response to the call for the use of assessment to promote greater student learning expressed in the National Higher Education Strategic Plan (2007-2020), and the increasing importance being attached to AfL at the international level. In the first set of modules, much attention is devoted to AfL elements such as formative assessment, giving quality feedback, and feedforward. Later modules explain different alternative assessment methods, as well as the procedures of developing valid and reliable traditional test formats with specific hands-on training on how to establish a test's validity and reliability. In sum, AKEPT's modules constitute a national attempt to educate Malaysian university lecturers on student learning assessment with the aim of striking a balance between AOL and AfL practices. Currently, efforts at AKEPT and individual university levels are ongoing to ensure that lecturers are properly trained in both.

Purpose and rationale of the study

In this article, we report a national study commissioned by AKEPT and the Malaysian Ministry of Higher Education to explore university lecturers' assessment practices and competencies *prior* to the development of the learning assessment modules. The study had set the stage for AKEPT's module development and assessment training efforts. The ultimate goal of the study was to establish whether lecturers' reported practices in assessment showed evidence of AfL, and if they would need training in its use. The study's specific objectives were:

1. To establish whether university lecturers' self-reported assessment practices constituted meaningful and interpretable dimensions of AfL;
2. To determine the relative frequency of AfL practices among the lecturers; and
3. To determine their perceived competencies in relevant aspects of AfL.

Research on Malaysian university lecturers' assessment practices is scarce. Prior to this study, there has been none that examined their assessment practices on a nationwide scale, particularly those related to AfL. The few works done thus far have explored lecturers' assessment practices within a single setting with a relatively small number of respondents (e.g. Zubairi, Sarudin and Nordin, 2008; Chan and Sidhu, 2013). This study was in part an attempt to address this gap in the literature.

Conceptual framework

The study was guided by the assessment competency framework developed by the American Federation of Teachers, the National Education Association and the National Council for Measurement in Education (1990). The framework consists of the following six standards for building teachers' competence in assessing student learning:

1. Choosing and developing assessment methods appropriate for instructional decisions;
2. Administering, scoring and interpreting the results of both externally-produced and teacher-produced assessment methods;
3. Using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement;
4. Developing valid grading procedures which use student assessments;
5. Communicating assessment results to students, parents, other lay audiences, and other educators; and

6. Recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.

The six standards can be used as a competency framework upon which quality assessment practices can be identified, implemented and accredited, as well as a guide to research on assessment. The study used five of the six standards as its conceptual foundation. The second competency (*i.e., administering, scoring and interpreting the results of both externally-produced and teacher-produced assessment methods*) was excluded as it was not relevant to the Malaysian higher education context. The use of the standards as a conceptual framework enabled the study to establish the multidimensionality of university lecturers' assessment practices, an area that is scarcely researched.

Literature Review

Barlowet *al.* (n.d.) reported that university lecturers were appreciative of the impact of AfL, finding it capable of changing instructional practices in positive ways. They found AfL effective in transforming the teaching and learning culture from one that was teacher-centered to that of a continuous student-teacher dialogue that drove instruction and planning. AfL also stimulated instructional behaviors that were more responsive to students' learning needs, and empowered them through ownership of learning. Similarly, Chan and Sidhu (2013) showed that both students and lecturers believed in the potential of formative assessment and feedback to transform the didactic learning in Malaysian universities. Students were in favour of formative assessment which they felt could lead to 'transformative learning due to the frequent and on-going constructive feedback they obtained from their lecturers' (Chan and Sidhu, 2013:6). Some highlighted that formative assessment activities encouraged critical reflection and allowed them to work at their own pace and ability, while others claimed that the activities helped them to make connections between theory and practice. The empirical support for these findings is documented in Black and Wiliam's (1998) review on teachers' use of AfL, in which the authors concluded that across educational levels and students' age groups, AfL had the biggest substantial impact on learning compared to other educational interventions with effect sizes ranging between $ES = .40$ and $ES = .70$.

Hake (2006) suggests it is high time that university lecturers' in Malaysia shift the higher education paradigm from teaching to learning. Thus, lecturers' adoption of AfL is a relevant agenda in the Malaysian higher education context. The first step in promoting a shift towards AfL among lecturers is to first understand how they view and practise assessment. However, much of the research on AfL has been done in the school context. Peterson and Einarson (2001: 629-630) pointed out that research into student assessment in higher education is still at its infancy, as there has been little 'empirical evidence concerning how institutions have conducted student assessment and to what effect'. In the Malaysian higher education context, research into assessment of student learning is scarce.

What we know about teaching staff's assessment competencies and practices is primarily drawn from research involving school teachers; and the limited literature in this area informs us that they are poorly prepared to practise AfL adequately (Dorn, 2010; Freeman and Lewis, 1998; Kibreab, 2011; Lewin, 2004; Mukki, 2012; Palomba and Banta, 1999; Sahari, 1999; Song and Koh, 2010; Townsend, 2007). According to Dorn (2010:328), 'although formative assessment is

appealing in theory, its practice as well as its definition is inconsistent'. This suggests that teachers and lecturers may lack an understanding of what AfL is due to inconsistencies that surround the issue. Mukki (2012) found that university lecturers' difficulty in practising AfL was reliably associated with insufficient training and exposure. More surprisingly, evidence from different countries revealed that even teacher educators themselves neither practised nor provided training in AfL adequately for them to act as promoters and role models of AfL (Lewin, 2004; Morris, 1996; Townsend, 2007; Zubairi, Sarudin and Nordin, 2008).

Studies that specifically examined teaching staff's competence in student assessment highlight that they do require training in AfL (Frazeir, 2007; Mukki, 2012; Zhang, 1995; Zhang and Burry-Stock, 1997; Zhang and Burry-Stock, 2003; Schaff, 2006). However, the results are mixed and point to different AfL aspects or competencies that they would require training in. For example, Zhang and Burry-Stock (1997) found Competency 3 – *'using assessment results when making decisions'* – to be the most difficult for teaching staff to practise, and Competency 5 – *'communicating assessment results'* – the least difficult. Contrary to Zhang and Burry-Stock (1997), Schaff (2006) discovered that *'using assessment results when making decisions'* was the least difficult. Instead, the staff surveyed reported that they had the most difficulty in *'developing valid grading procedures'* (Competency 4). In another study, the results indicated that *'developing assessment methods'* (Competency 1) was a difficult experience (Frazeir, 2007).

Several reasons may be used to explain the mixed findings. A plausible one is a lack of clarity regarding what the indicators of each standard mean, and this makes them open to wide and conflicting interpretation. It is also very likely that teachers and lecturers do not share a common understanding of the meaning and requirements of AfL (Dorn, 2010; Dunn and Mulvenon, 2009). Hence, they could have drawn the meaning of the indicators based on whatever they know about assessment, rather than on deep-seated knowledge of it. Therefore consistent findings regarding AfL practices, particularly among higher education lecturers, could not be replicated across studies.

Methodology

Sample

The sample consisted of 534 lecturers from 33 public and private higher learning institutions in Malaysia. The male and female representation in the sample was adequate, with male lecturers making up about 52.7% of the composition. Close to half (48.8%) were professors, associate professors and senior lecturers; the rest were assistant lecturers, tutors and other categories of teaching staff. A majority (51.7%) were teaching in the areas of applied sciences and technology. The duration of teaching experience among the respondents ranged between one and 44 years with a mean experience of 10 years. Table 1 summarizes the sample's characteristics.

Table 1. Percentage and Frequency Distribution of Sample Characteristics (N=534).

Sample Characteristics	%	Frequency
Gender		
Male	52.7	281
Female	47.3	253
Position		
Professor	4.7	25
Associate Professor	13.3	71
Senior Lecturer	30.9	165
Lecturer & Other Academic Positions	51.2	273
Highest Qualification		
Doctoral Degree	45.6	243
Master's Degree	45.6	243
Bachelor's Degree	8.8	48
Discipline of Study		
Pure Sciences	4.6	25
Applied Sciences & Technology	51.7	276
Medical Sciences	7.6	40
Social Science & Humanities	36.1	193
Institution Type		
Public	43.0	230
Private	57.0	304

Instrument

A 24-item assessment practice inventory was used to measure the two constructs of interest, namely (i) lecturers' self-reported practices in AfL (which would enable the analysis to extract meaningful and interpretable dimensions of AfL practices from the data, as well as frequency of practice), and (ii) their perceived competency in the given AfL aspect. The inventory measured practices on a 10-point Likert scale ranging from 1 (Never) to 10 (Always), and competency also on a 10-point scale of 1 (Not At All Competent) to 10 (Highly Competent). The items were drawn from Mukki's (2012) Assessment Practices Inventory for Teacher Education (APITE), which he revised from the original Assessment Practices Inventory (API) developed by Zhang (1995), and Zhang and Burry-Stock (2003). The 24 items represented five of the six standards stipulated in the assessment competency framework developed by the American Federation of Teachers, the National Education Association and the National Council for Measurement in Education (1990). The study made use of expert-judgment to content-validate the importance and relevance of the items prior to data collection.

Access

Access to the respondents was obtained through a letter sent by AKEPT to the Vice President (Academic Affairs) of each of the 33 higher education institutions in Malaysia. AKEPT's letter explained the nature and purpose of the study, and related its importance to the achievement

of quality assessment in the country's higher learning institutions. The Vice President's office then wrote to all deans and faculties to urge lecturers to participate in the survey. The directive to participate was, therefore, top-down.

Results

Dimensions of AfL Extracted from Malaysian University Lecturers' Self-Reported Assessment Practices

To address the first research objective, the data were subjected to the procedures of Principal Component Analysis (PCA). The purpose of this analysis was to extract underlying dimensions of assessment practice from the data, and ascertain whether these dimensions constituted meaningful and interpretable practices of AfL. Hence the PCA procedures would allow the study to determine lecturers' adherence to the five stipulated assessment standards.

First the data were checked for sampling adequacy. In PCA, this is indicated by the Kaiser-Meyer-Olkin (KMO) measure where a value of .60 is the suggested minimum to show an adequate sample size. According to Kaiser (1974), any value exceeding 0.8 is considered meritorious in terms of sampling adequacy. For this study, the KMO measure was .86 which indicates that the sample size requirement was well met. The degree of intercorrelation among the items also justified the use of PCA. Bartlett Sphericity Test was statistically significant, $\chi^2(105) = 4310, p < .001$. These indices showed that the sample size and data were fit for the application of PCA procedures.

The PCA analysis revealed that out of the inventory's 24 items, 14 were found to represent meaningful and interpretable dimensions of AfL. The remaining 10 failed to load into any dimension, suggesting that the items did not represent meaningful assessment practices among the respondents. Table 2 shows the correlation matrix and communality for the 14 items. The correlation values in the matrix indicate how closely related each item or practice is to one another.

Table 2.Correlation Matrix and Communality.

	S1a	S1b	S1c	S1d	S1e	S4a	S4c	S4d	S3b	S3c	S3d	S2a	S2b	S2c
S1b	.567													
S1c	.494	.554												
S1d	.427	.404	.667											
S1e	.409	.506	.588	.579										
S4a	.328	.416	.473	.433	.398									
S4b	.414	.391	.452	.432	.395	.781								
S4d	.343	.410	.496	.478	.422	.762	.705							
S3b	.375	.399	.364	.287	.276	.269	.388	.334						
S3c	.272	.275	.336	.270	.238	.239	.280	.254	.495					
S3d	.283	.292	.280	.260	.240	.168	.272	.185	.620	.537				
S2a	.280	.303	.301	.243	.284	.272	.295	.258	.318	.232	.356			
S2b	.217	.251	.223	.173	.192	.191	.239	.202	.283	.237	.263	.426		
S2c	.206	.264	.223	.157	.194	.198	.216	.216	.262	.265	.256	.407	.920	
Com	.58	.60	.71	.62	.63	.80	.77	.84	.68	.62	.74	.42	.93	.92

To extract a factor solution, the analysis used the correlated-factor method of axis rotation. This method was chosen because the AfL dimensions to be extracted from the data were assumed to be theoretically correlated. Table 3 reports the rotated pattern matrices of the analysis.

Table 3. Loadings for Four-Factor (Correlated-Factor) Solution

Practice	Code	Dimension			
		S1	S2	S3	S4
Evaluating oral questions from students	S1a	.77	.02	.10	-.14
Assessing student learning through observations	S1b	.75	.08	.06	-.04
Assessing individual students' class participation	S1c	.79	-.01	.01	.10
Assessing group participation	S1d	.70	-.07	-.03	.19
Assessing individual students' hands-on activities	S1e	.81	.00	-.09	.04
Incorporating language ability in the calculation of grades	S4a	.06	.02	-.05	.87
Incorporating effort in the calculation of grades	S4c	-.06	.03	.14	.85
Incorporating student attentiveness in the calculation of grades	S4d	.06	.01	-.01	.88
Providing oral feedback to students	S3b	.05	.00	.79	.09
Providing written feedback to students	S3c	-.01	-.02	.78	.05
Communicating assessment results to students	S3d	-.01	.02	.89	-.07
Communicating assessment results to other educators	S2a	.12	.49	.16	.06
Recognising unethical, illegal, or inappropriate assessment methods	S2b	-.03	.98	-.04	.00
Recognising unethical, illegal, or inappropriate uses of assessment information	S2c	-.03	.98	-.05	.00
Internal Consistency (Cronbach's Alpha)		.84	.82	.78	.88

Note: S1 –Using diverse AfL methods;
 S2 –Recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information;
 S3 –Communicating assessment results to students
 S4 –Grading practices

Four underlying dimensions of assessment practice compliant with AfL practices were extracted from the data. The reproduced correlation matrix seemed to best fit the intercorrelation among the 14 distinct assessment practices, accounting for 70% of the total variance. The first dimension explained 40% of the total variance; it extracted the largest eigenvalue (6.0), while the subsequent eigenvalues were 1.9, 1.4, and 1.1 respectively. The final solution was free from factorial complexity and item-specific factor. These results enhanced our comfort that we were dealing with non-chance loadings. What this means is that the same results may be replicable across different university settings.

The four-factor solution indicated that the first five items (i.e. *evaluating oral questions from students, assessing student learning through observations, assessing individual students' class participation, assessing group participation, and assessing individual students' hands-on activities*) loaded on the first dimension (S1). Collectively these items represent the use of diverse AfL methods. Their empirical clustering perfectly matches the logical and theoretical

grouping of the items. These items measure the degree to which a lecturer practises the different methods of AfL. High scores on this dimension imply that the lecturer practises assessment methods appropriate for instructional decisions, which indicates adherence to the first standard for teacher competence in assessment.

The second rotated factor (S2) was represented by three items, namely *'communicating assessment results to other educators'*, *'recognising unethical, illegal, or inappropriate assessment methods'*, and *'recognizing unethical, illegal, or inappropriate uses of assessment information'*. This factor indicates the degree to which a lecturer complies with the sixth competence standard, which is *recognizing assessment methods and uses of assessment information that are unethical, illegal, and otherwise inappropriate in nature*. Items on *'communicating results to students'*, *'providing oral feedback'*, and *'providing written feedback'* were found to load substantially on the third factor (S3), which represents the standard for communicating assessment results and feedback. The final dimension (S4) extracted from the analysis also contained three items (*i.e. incorporating language ability, effort and student attentiveness in the calculation of grades*) with strong and significant loadings (all of which were above .80). The items collectively represent lecturers' grading practices.

The results provided evidence of credible lecturer practices that adhered to four of the five standards of assessment competency used as the study's conceptual foundation. However, ten of the inventory's 24 items failed to load on any of the extracted dimensions. These theoretically important indicators of AfL, for example *'using assessment results for instructional decisions'*, *'aligning assessment with instruction'*, *'developing grading philosophy'*, *'using assessment results when evaluating class improvement'*, and *'using portfolios to assess student progress'* cross loaded on more than one dimension of the assessment competency standards. This suggests that these items did not constitute common or meaningful practices among the lecturers.

Frequency of AfL Practices among Malaysian University Lecturers

The study examined the descriptive statistics drawn from the PCA results, which constituted the mean score and standard deviation of each of the AfL practices and its indicators. The resulting mean scores for these constructs and items enabled the study to rank-order and compare the relative frequency of AfL practices among lecturers. Table 4 summarises the distribution of mean scores according to the extracted constructs and their indicators.

The mean scores of individual items appeared to range between 5.85 (*'incorporating language ability in the calculation of grades'*) and 8.20 (*'communicating assessment results to students'*). The value of 10 indicates that the given act is always practised, while the value of 1 indicates that it is never practised. Based on the individual mean scores, it can be concluded that Malaysian university lecturers quite frequently communicated assessment results to students (M=8.20), gave oral feedback (M=8.12), assessed group participation (M=7.64), and used observations to assess learning (M=7.57). They also quite frequently assessed students' class participation (M=7.43), hands-on activities (M=7.42), and oral questions (M=7.41).

Table 4. Mean Ranking of Assessment Practices (Maximum Score = 10 Indicating Always Being Practised).

	Mean	Std. Deviation	Construct Mean
Communicating Results and Feedback to Students			7.88
Communicating assessment results to students	8.20	1.65	
Providing oral feedback to students	8.12	1.67	
Providing written feedback to students	7.31	2.06	
Using Diverse AfL Methods			7.49
Assessing group participation	7.64	1.97	
Assessing student learning through observation	7.57	1.93	
Assessing individual student's class participation	7.43	2.00	
Assessing individual student's hands-on activities	7.42	2.01	
Evaluating oral questions from students	7.41	2.10	
Recognising Unethical, Illegal, and Inappropriate Assessment			6.84
Communicating assessment results to other educators	6.90	2.28	
Recognizing unethical, illegal, or inappropriate assessment methods	6.87	2.68	
Recognizing unethical, illegal, or inappropriate uses of assessment information	6.76	2.68	
Grading Practices			6.15
Incorporating effort in the calculation of grades	6.36	2.45	
Incorporating student attentiveness in the calculation of grades	6.24	2.52	
Incorporating language ability in the calculation of grades	5.85	2.61	

Based on the construct mean scores, the AfL aspects most frequently practised were '*communicating results and feedback to students*' (M=7.88) and '*using diverse AfL methods*' (M=7.49). The other two AfL aspects were practised to a lesser extent by Malaysian lecturers.

Perceived Competency in AfL Practices Among Malaysian University Lecturers

Table 5. presents Malaysian lecturers' competency level in AfL practices.

Table 5. Mean Ranking of Assessment Competency (Maximum Score = 10 Indicating Highly Competent).

	Mean	Std. Deviation	Construct Mean
Communicating Results and Feedback to Students			7.7
Communicating assessment results to students	8.0	2.1	
Providing oral feedback to students	8.0	2.1	
Providing written feedback to students	7.3	2.3	
Using Diverse AfL Methods			7.3
Assessing group participation	7.4	2.3	
Assessing student learning through observation	7.3	2.2	
Assessing individual student's class participation	7.4	2.0	
Assessing individual student's hands-on activities	7.4	2.0	
Evaluating oral questions from students	7.4	2.1	
Recognizing Unethical, Illegal, and Inappropriate Assessment			6.8
Communicating assessment results to other educators	6.9	2.4	
Recognizing unethical, illegal, or inappropriate assessment methods	6.7	2.7	
Recognizing unethical, illegal, or inappropriate uses of assessment information	6.7	2.7	
Grading Practices			6.3
Incorporating effort in the calculation of grades	6.6	2.4	
Incorporating student attentiveness in the calculation of grades	6.4	2.6	
Incorporating language ability in the calculation of grades	6.0	2.6	

A score approaching 10 indicates high competence in the given practice. The reported competency levels ranged between 6.0 and 8.0. Based on the mean scores of individual items, two items stood out as practices that lecturers were most competent in. They are '*communicating assessment results to students*' and '*providing oral feedback to students*', both with a mean score of 8.0. Lecturers reported a lower competency level for '*providing written feedback*' (M=7.3) and using diverse AfL methods (mean scores ranging between 7.3 and 7.4). Grading practices that take into account students' effort, attentiveness and language ability were AfL aspects that lecturers reported to be the least competent in.

Based on the construct means, the AfL dimension that lecturers perceived to be quite competent in was '*communicating results and feedback to students*', and '*using diverse AfL methods*'. On the other hand, the competency level reported for '*grading practices*' was relatively low (M=6.3). The same could be said about the reported competency level for recognizing unethical uses of assessment information (M=6.8). Apparently, the rank order of these AfL competency dimensions was identical to that of AfL practices.

Discussion and Conclusion

Several findings are worthy of highlight. Firstly, the data supported the multidimensionality of AfL practices. The extracted dimensions adhered to four competency standards stipulated in the assessment competency framework developed by the American Federation of Teachers, the National Education Association, and the National Council for Measurement in Education (1990). The findings replicated the results of previous research on AfL (Frazee, 2007; Mukki, 2012; Zhang, 1995; Zhang and Burry-Stock, 1997; Zhang and Burry-Stock, 2003; Schaff, 2006), which also found it to be a multidimensional construct.

Secondly, in terms of practice, the results suggested that Malaysian university lecturers do use AfL but their practices were limited to four underlying dimensions, namely (i) communicating assessment results and feedback, (ii) using diverse AfL methods, (iii) recognising unethical, illegal, and inappropriate assessment methods, and (iv) employing grading practices that incorporated students' effort, attentiveness and language ability. Among the four underlying dimensions, the first two were found to be quite common among the lecturers, while the last two were practised to a lesser extent. The finding is consistent with the results of Mukki (2012) and Zhang and Burry-Stock (1997). Both found that 'communicating assessment results' was the least difficult, while 'using assessment results when making decisions' was the most difficult standard to be adhered to by teaching staff. In addition, Schaff (2006) also reported 'developing grading procedures' to be the most difficult standard for higher education teachers.

Thirdly, the data did not provide evidence that Malaysian university lecturers practised several other critical aspects of AfL such as conducting formative assessment and giving feedforward to guide further learning. From what we know about their long-standing traditional assessment practices, it is likely that they were unfamiliar with formative assessment and what feedforward looks like, and would thus need greater exposure to and specific training in these two aspects. The results offer sufficient evidence to support the development of an intervention programme that should aim at enhancing Malaysian university lecturers' understanding of AfL and its fundamental practices.

Fourthly, with respect to perceived competency in AfL, the lecturers reported being most capable in communicating assessment results and feedback. This was followed by competencies in using diverse AfL methods, recognising unethical, illegal, and inappropriate assessment methods, and employing grading practices that incorporated students' effort, attentiveness and language ability. It appeared that the ordering of frequency of practice was identical to that of perceived competence, suggesting a possible reliable association between lecturers' assessment practice and their perceived competence in AfL. The pattern of similarity suggests that these two constructs may be positively correlated. In other words, higher levels of perceived competence would likely lead to more frequent AfL practices, or vice versa. This observation calls for intervention programmes that could enhance university teachers' competence in AfL. The postulated causal relationship between practice and competency also merits further examination in future studies.

Its contributions notwithstanding, the study was not without limitations. The most important one concerns the limitation of the assessment inventory itself – that it does not adequately

account for several fundamental indicators of AfL, such as formative assessment, constructive alignment of assessment practices, personal reflection, and use of portfolios and technology-based methods to support AfL (Barkhi & Williams, 2010; Heinrich *et al.*, 2009; de Jesus and Moreira, 2009; Liang and Creasy, 2004; Meeus, Petegem, and Engels, 2009). Nor does the inventory include indicators related to lecturers' evidence-based practice or the nine principles of good practice for assessing student learning (Astinet *et al.*, 1996). This concern may be addressed in future research examining teachers' and university lecturers' use of AfL.

The inventory's limitation in broadly measuring AfL gives rise to another issue of concern, which is the efficacy of the reported practices in explaining student learning. For example, some may question how lecturers' frequency of communicating assessment results and feedback to students contributes to the improvement of student learning. Quite recently Price *et al.* (2010:277) argued against 'many of the assumptions and beliefs about the effectiveness of feedback practices'. It has hence been questioned whether feedback matters in student learning. We believe that it certainly does, especially when the mediators, suppressors and moderators of the relationship between feedback and student learning could be identified and addressed accordingly. However, this issue warrants a systematic examination, the results of which may enable the inventory to be improved.

In sum, the study has added valuable insights into further understanding university lecturers' AfL practices in the classroom despite its limitations. The data generated are useful in informing ongoing efforts in designing and implementing intervention programmes to promote and support AfL adoption among higher education lecturers in Malaysia. The study concurs with the view of Black and Wiliam (1998:2-3) that:

'...fundamental change in education can be achieved only slowly -- through programs of professional development that build on existing good practice. Thus we do not conclude that formative assessment is yet another "magic bullet" for education. The issues involved are too complex and too closely linked to both the difficulties of classroom practice and the beliefs that drive public policy.'

In our context, we believe that AfL is one of the promising means of transforming the traditionalist, teacher-oriented paradigm of Malaysian higher education into one that is more learning-centered, as suggested by Hake (2006). This goal may partly be achieved through the activities organized by AKEPT and the proactive role played by the Malaysian Ministry of Higher Education as the primary change agent towards this end. To facilitate this transformation process, the issues and challenges that confound AfL implementation in Malaysian universities should be immediately identified and addressed in relevant research and initiatives.

Acknowledgement

The study was conducted with the funding and managerial assistance from the Malaysian Higher Education Leadership Academy (AKEPT). The research team would like to thank in particular the Director of the academy, Professor Dr. Majid Konting, for his trust in the team and continuous moral support.

References

- American Federation of Teachers, National Council on Measurement in Education, and National Education Association (AFT, NCME, & NEA) (1990) *Standards for teacher competence in educational assessment of students*. Washington, DC: National Council on Measurement in Education.
- Anderson, T. (2004) Teaching in an online learning context, in Anderson, T.&Elloumi, F. (eds.), *Theory and practice of online learning*. Canada: Athabasca University,273-294.
- Angelo, T.A. (1995) Reassessing (and defining) assessment, *AAHE Bulletin*, 48(3), 7.
- Astin, A.W., Banta, T.W., Cross, K.P., Khawas, E., Ewell, P.T, Hutchings, P.T., Marchese, T.J., McClenney, K.M., Mentkowski, M., Miller, M.A., Moran, E.T., and Wright , B.D. (1996) 9 principles of good practice for assessing student learning. *AAHE Assessment Forum*. Available at:
http://www.academicprograms.calpoly.edu/pdfs/assess/nine_principles_good_practice.pdf(Accessed: 5 March 2012).
- Australian National Training Authority (ANTA). (2002)*Assessment and online teaching: Australian flexible learning quick guide series*. Available at:
<http://www.flexiblelearning.net.au/guides/assessment.pdf> (Accessed: 8 February 2012).
- Barkhi, R., and Williams, P. (2010)The impact of electronic media on faculty evaluation,*Assessment and Evaluation in Higher Education*, 35(2), 241-262.
- Barlow, S., Bertrand, M.,Majkot, V.S., McLaughlin, K., and Speir, S. (n.d.)*Action research projects focused on the impact of assessment for learning strategies on teacher practice and student learning*. Available at:
http://www.rainbowschools.ca/community/support/Overview_of_Action_Research_AfL_2009-10.pdf.(Accessed: 15 January 2012).
- Black, P., and Wiliam, D. (1998) Inside the black box: Raising standards through classroom assessment.*Phi Delta Kappan*. Available at:
http://blog.discoveryeducation.com/assessment/files/2009/02/blackbox_article.pdf.(Accessed: 15 July 2012).
- Black, P., Harrison, C., Lee, C., Marshall, B., and Wiliam, D. (2004) Working inside the black box: assessment for learning in the classroom,*Phi Delta Kappan*, 86(1), 9–21.
- Boud, D. (1995) *Enhancing learning through self assessment*. London: Kogan Page.
- Brown, G., Bull, J., and Pendlebury, M. (1997)*Assessing student learning in higher education*. London: Routledge.
- Chan Y.F., and Sidhu, G.K. (2013) Promoting transformative learning through formative assessment in higher education,*ASEAN Journal of Teaching and Learning in Higher Education*, 5(1), 1-11.
- de Jesus, H.D., and Moreira, A.C. (2009) The role of students' questions in aligning teaching, learning and assessment: a case study from undergraduate sciences, *Assessment and Evaluation in Higher Education*, 34(2), 193-208.
- Dorn, S. (2010)The political dilemmas of formative assessment, *Exceptional Children*. 76(3), 325-337.
- Dunn, K.E., and Mulvenon, W. (2009)A critical review of research on formative assessment: the limited scientific evidence of the impact of formative assessment in education,*Practical Assessment and Evaluation*.14(7). Available at:
<http://www.http://pareonline.net/genpare.asp?wh=0&abt=14>.(Accessed: 20 March 2012).

TUNKU AHMAD, ZUBAIRI, IBRAHIM, OTHMAN, RAHMAN, RAHMAN, NORDIN& NOR:
ASSESSMENT FOR LEARNING PRACTICES AND COMPETENCY AMONG MALAYSIAN UNIVERSITY
LECTURERS: A NATIONAL STUDY

- Ellyn, G. (2000) *Assessment: an institution-wide process to improve and support student learning*. College of DuPage: ERIC Reproduction Services: ED 450 833.
- Ewell, P. T. (2009) *Assessment, accountability, and improvement: revisiting the tension*, NILOA Occasional Paper No.1, Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment. Available at: http://www.learningoutcomeassessment.org/documents/PeterEwell_005.pdf (Accessed: 20 March 2012).
- Frazeir, C.H. (2007) *Investigating teachers' self-perceived application of classroom assessment practices using the assessment practices inventory revised (APIR)*. Unpublished doctoral dissertation. University of Alabama.
- Freeman, R., and Lewis, R. (1998) *Planning and implementing assessment*. London: Kogan Page.
- Hake, R.R. (2006) Measuring teaching and learning performance: interconnected issues, *the Third International Conference on Measurement and Evaluation in Education (ICMEE)*. Penang, Malaysia, 13-15 February. Available at: <http://www.physics.indiana.edu/~hake/ICMEEK-2006.pdf> (Accessed: 16 November 2012).
- Hannafin, M., Hill, J. R., Oliver, K., and Glazer, E. (2003) Cognitive and learning factors in web-based distance learning environments, in: Moore, M. G. and Anderson, W. G. (eds.) *Handbook of distance education* Mahwah, NJ: Lawrence Erlbaum Associates.
- Harlen, W., and Crick, R.D. (2003) Testing and motivation for learning, *Assessment in Education*, 10(2), 169-207.
- Heinrich, E., Milne, J., Ramsay, A., and Morrison, D. (2009) Recommendations for the use of e-tools for improvements around assignment marking quality, *Assessment and Evaluation in Higher Education*, 34(4), 469-479.
- Kaiser, H.F. (1974) An index of factorial simplicity, *Psychometrika*, 39, 31-36.
- Kibreab, A.E. (2011) Instructors' understanding and practice of continuous assessment: the case of ethiopian civil service college, *Journal of Quality and Relevant Education and Training*, 1(1), 33-42.
- Klenowski, V. (2009) Assessment for learning revisited: an asia-pacific perspective, *Assessment in Education: Principles, Policy and Practice*, 16(3), 263-268.
- Lewin, K.M. (2004) *The pre-service training of teachers: does it meet the objectives and how it could be improved?* Paper commissioned for the EFA Global Monitoring Report 2005, The Quality Imperative. Available at: <http://unesdoc.unesco.org/images/0014/001466/146658e.pdf> (Accessed: 18 August 2012).
- Liang, X., and Creasy, K. (2004) Classroom assessment in web-based instructional environment: instructors' experience, *Practical Assessment, Research and Evaluation*, 9(7).
- Marzano, R. J. (2000) *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Meeus, W., Petegem, P.V., and Engels, N. (2009) Validity and reliability of portfolio assessment in pre-service teacher education, *Assessment and Evaluation in Higher Education*, 34(4), 401-413.
- Morris, J.B. (1996) *Assessment instruction in preservice teacher education courses at University of South Carolina*. Unpublished doctoral dissertation. University of South Carolina.
- Mukki, R.A. (2012) *A study of assessment practices in teacher education programmes in Tanzania*. Unpublished doctoral thesis. International Islamic University Malaysia.

TUNKU AHMAD, ZUBAIRI, IBRAHIM, OTHMAN, RAHMAN, RAHMAN, NORDIN & NOR:
ASSESSMENT FOR LEARNING PRACTICES AND COMPETENCY AMONG MALAYSIAN UNIVERSITY
LECTURERS: A NATIONAL STUDY

- Ong, S.L. (n.d.) *Profiling Classroom Teachers Assessment Practice*. Unpublished research paper. Available at: http://www.iaea.info/documents/paper_4d32f2cd.pdf (Accessed: 20 November 2012).
- Palomba, C.A., and Banta, T.W. (1999) *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco: Jossey-Bass.
- Peterson, M.W., and Einarson, M.K. (2001) What are colleges doing about student assessment? *The Journal of Higher Education*, 72(6), 629-669.
- Price, M., Handley, K., Millar, J., and O'Donovan, B. (2010) Feedback: all that effort, but what is the effect? *Assessment and Evaluation in Higher Education*. 35(3), 277-289.
- Riley, K., and Stern, B. (1998) Using authentic assessment and qualitative methodology to bridge theory and practice, *Educational Forum*, 62(2), 178-185.
- Sahari, M. (1999) The perceptions of university instructors on their testing practices: A case study, *Pertanika Journal of Social Sciences*, 7(1), 21-29.
- Schaff, C.S. (2006) *A Rasch model analysis of standards for teacher competence in educational assessment of students using the classroom assessment inventory*. Unpublished doctoral dissertation. The Northern Illinois University.
- Song, E., and Koh, K. (2010) Assessment for learning: understanding teachers' beliefs and practices. *The 36th Annual Conference Assessment for the Future Generations*, Bangkok, Thailand, 20 – 27 August.
- Stiggins, R. (2002) Assessment crisis: the absence of assessment for learning, *Phi Delta Kappan*. 83(10), 758-765.
- The Quality Assurance Agency for Higher Education. (2011) *Understanding assessment- its role in safeguarding academic standards and quality in higher education: a guide for early career staff*. Available at: <http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/UnderstandingAssessment.pdf> (Accessed: 15 January 2014).
- Townsend, A.E. (2007) *Thai university teacher education programmes reform: professor perspectives of pre-service teacher preparation in assessment of student learning*. Unpublished doctoral dissertation. Washington State University.
- Zhang, Z. (1995) *Investigating teachers' self-perceived assessment practices and assessment competencies on the assessment practices inventory*. Unpublished doctoral dissertation. Tuscaloosa, AL: The University of Alabama.
- Zhang, Z., and Burry-Stock, J.A. (1997, March) Assessment practices inventory: a multivariate analysis of teachers' perceived assessment competence. The Annual Meeting of the National Council of Measurement in Education, 25-27 March. Chicago.
- Zhang, Z., and Burry-Stock, J.A. (2003) Classroom assessment practices and teachers' self-perceived assessment skills, *Applied Measurement in Education*, 16(4), 323-342.
- Zubairi, A.M., Sarudin, I., and Nordin, M.S. (2008) Faculty competency in assessment. *The Annual Conference of the Association for Educational Assessment (IAEA)*, 7-12 September. UK: Cambridge.