

Cambridge Horizons, 3 June 2013

School-based Assessment: Prospects and Realities in Asian Contexts

This document is a compilation of papers provided by the speakers at the Cambridge Horizons seminar on **School-based Assessment: Prospects and Realities in Asian Contexts**.

The seminar was held in Subang Jaya Malaysia on 3 June 2013.

The speakers and the title of their presentations at the seminars are:

High Stakes School-Based Assessment and Cultural Values: Beyond Issues of Validity

By Prof Kerry Kennedy
Dean, Faculty of Education and Human Development
Hong Kong Institute of Education

School-Based Assessment, Standards, Teacher Judgement and Moderation: Messages from Research

By Prof Val Klenowski
Faculty of Education
School of Cultural and Professional Learning
Queensland University of Technology, Australia

School-based Assessment: Transformation in Educational Assessment in Malaysia

By Hjh Norzila Mohd Yusof
Examinations Syndicate, Ministry of Education Malaysia

Additional resources from the seminar will be available here:

www.cambridgeassessment.org.uk
www.cambridgeassessment.org.uk/singapore

High Stakes School Based Assessment and Cultural Values : Beyond Issues of Validity¹

Kerry J Kennedy²

Abstract

Assessment plays a key role in all societies but a very special role in Asia. A firm commitment to meritocracy permeates community thinking and values in the region. This commitment is coupled with a belief in education as the process by which talent will be recognized irrespective of social class, cultural capital and family networks. Testing and examinations are trusted to secure this important social end.

It is in this cultural context that school based assessment (SBA) must be appreciated. Promoted in different parts of the region and beyond the rationale is most often linked to validity. The technical rational for SBA is relatively easy to develop and promote – more valid assessments should increase both the fairness of the testing process and public confidence in that process. A side effect should be a more professional teacher work force taking responsibility for important educational and social processes.

Yet the introduction of SBA in Asian societies has not always been so straightforward. Increased teacher and student workload, lack of community confidence in school based processes and even lack of confidence by teachers themselves emerge as key issues during implementation. Perhaps of more significance is the reluctance to accept responsibility for high stakes school based assessment. Thus SBA is often at risk not because it is a poor assessment process but because the conditions in which it is implemented are not always conducive. This paper will consider these conditions from the point of view of policy makers, teachers and the community.

Key Note Address, Cambridge Horizon's Seminar, 'School based assessment: Prospects and realities in Asian contexts', 3 June 2013, Kuala Lumpur, Malaysia.

¹ Part of this paper draws on "Integrating Assessment of Learning and Assessment for Learning in Hong Kong Public Examinations: Rationales and Realities of Introducing School-based Assessment", Technical Paper No. 3 of the QEF Funded Project, *Assessment for Productive Learning: Forms of Assessment and their Potential for Enhancing Learning*.

² Professor Kerry Kennedy is Chair Professor of Curriculum Studies at the Hong Kong Institute of Education where he also holds concurrent appointments as Associate Vice-President (Quality Assurance), Dean of the Faculty of Education and Human Development and Director of the Centre for Governance and Citizenship

Assessment is a ubiquitous process that characterizes much of everyday life: we make judgments about which clothes, perfume and jewelry we like; which paintings we prefer and those we don't; the food that appeals and that which doesn't. These are all assessments of a kind, qualitative in nature, subjective and not always consistent. The importance of such assessment has been brought home to us in the work of Elliott W Eisner (1991) who developed his connoisseurship model of educational evaluation. From another perspective, when we visit the doctor, he/she makes instant judgments about whatever aches or pains we present and instantly prescribes this or that medication – often with the proviso that we should come back next week if the aches and pains persist. We often assume that doctors' judgments are scientific but Western medicine is much more intuitive – much more inductive than we would like to believe. Yet we all survive with the assessments of daily living – their tentativeness, their inconsistency and subjectivity.

But when it comes to schools we demand much more. We want valid and reliable assessments that are transparent and fair and that can be used for making decisions – life changing decisions in some cases. Assessment in educational contexts has a high stakes quality about it that does not always apply to the kind of assessment we make for our daily living. Assessment in education can determine the school students attend and the subjects they take. It can also determine students' progress from one class to the next, from one level of schooling to the next including from school to university and beyond. This is often called the 'sifting and sorting' function of assessment – what Yu, Kennedy, Fok & Chan (2009, p.8) called the "instrumental aspects" of assessment. Not all assessment is instrumental in nature, as the advocates of assessment for learning have reminded us (Assessment Reform Group, 1999), but it is important to keep in mind that the purposes of assessment are not always educational. This is particularly true when it comes to school based assessment (SBA) when it is not always easy to distinguish between its educational purposes and its instrumental purposes. This confusion over purposes is one of the major issues to confront policy makers, practitioners and the community and this will be the focus of the remainder of this paper.

In what follows, three key issues will be discussed to highlight issues connected to the purposes of SBA:

1. Identifying different forms of “school based assessment” and their relationship to “high stakes SBA”, the topic of this paper.
2. Locating high stakes SBA in the social, cultural and educational contexts of Asian societies;
3. Examining criteria for successful high stakes SBA in Asian contexts.

Assessment for learning and SBA – similarities and differences and why they are important

Classrooms as sites for assessment are by no means a new phenomena but when a high profile group such as OECD (2005) advocate for formative assessment then you know classroom assessment has to be important. Supported by the work of the Assessment Reform Group (1999) in the United Kingdom, the focus on providing feedback to students on the progress they are making towards learning goals has become almost a mantra for assessment specialists. Empirical research seems to suggest that formative assessment can make a real difference to student learning. In Hong Kong, the major education reform that started around 2001 highlighted the importance of reorienting assessment in this direction.

With a focus on “assessment for learning”, teachers in Hong Kong have been encouraged to view assessment not only as examinations and tests, but also as part of a learning process that can provide feedback to students to help them improve their learning (Curriculum Development Council, 2001). The Education Commission (2000) proposed to eliminate excessive dictation exercises, mechanical drills, tests and examinations and recommended the use of various modes of assessment including flexible formative assessment. The Curriculum Development Council (2001) suggested that:

- (a) teachers provide feedback to students of their strengths and weaknesses and;
- (b) schools include key attitudes, self-management, and moral and civic qualities in report cards as part of student achievement and as a basis for further improvement.

The Council made assessment *for* learning the prime target in all its proposed measures, which included:

1. evidence-based quality criteria in line with the curriculum framework;
2. combined curriculum and assessment guides for each subject to make assessment and objectives consistent; and
3. liaison with the universities about broadening university admission criteria.

At the same time as Hong Kong teachers were encouraged to focus more on classroom assessment, the old British style examination system was coming to an end to be replaced by a single end-of-school examination. Thus while assessment for learning was considered to be important – especially as a way of moderating the influences of poor teaching and bad learning habits- the end of Form 6 examination meant that teachers were not entirely free to utilize and take advantage of classroom assessment. Teaching for examinations, at least in the senior years of schooling, remained a challenge both for teachers and students. As Biggs (1996) pointed out, examinations and norm referenced assessment were very much part of Hong Kong's colonial experience – moving to assessment for learning was a big step and linking examinations to assessment for learning an even bigger step. Yet as Hong Kong's education system moved forward under the impetus of reform from 2001 onwards, it was this new approach to assessment that dominated reform discourse.

To pave the way for the new examination system SBA was trialled in the old system in relation to English oral assessment. Yet this was not Hong Kong's first experience with SBA. It had been implemented in a number of subjects like Advanced Level (AL) Chemistry and HKCEE Design and Technology and Electronics and Electricity since 1978 (HKEAA, 2005) as well as AL Biology and AL Physics School since 2003 (Yung, 2001, Yip & Cheung, 2005). SBA was referred to at that time as the Teacher Assessment Scheme and the responses from teachers were already known as Yip & Cheung (2005, p.161) pointed out:

Many teachers view teacher assessment as additional work imposed on them by the authorities and, together with the lack of implementation skills and supporting resources, the scheme adds extra workload and pressure to their routinely busy timetable.

This becomes a well worn theme whenever SBA is introduced but, it was not the only issue identified in Hong Kong prior to full scale implementation of the new examination and assessment system.

Cheung (2001) reported that in an open-ended survey (n=53) teachers expressed a wide range of worries, for example, their ability to design high-quality assessment tasks, whether SBA will affect teacher-student relationship, assessment method to be used etc. A strategic review of assessment identified an even more serious issue involving teachers in SBA the public does not trust teachers (IBM, 2003, p.34):

One reason put forward in the past for limiting school-based assessment is the public's ostensible lack of trust of teachers. This has been too readily accepted as received wisdom. It dates from a period before teaching was a trained profession, and it no longer fits the modern education system.

Numerous questions about the roles and judgments of teachers were raised, related to validity and reliability in SBA (Chang, 2004; Hau, 2004). In one sense they are technical questions but they go to the heart of public confidence in SBA and therefore they are not always easy to address (Chang, 2004). Broadfoot & Black (2004, p.16) rightly commented that teachers' role in summative assessment is not easy to be recognized.

Finally, the examination-oriented approach generally accepted in Hong Kong society has to be seen the biggest issue in SBA's implementation. This approach makes the society focus on assessment of learning. A high proportion of teachers and students emphasized on doing well in examinations and many are proud of their success (Cheng, 2004). This approach is not new. In 1982, a report that had an impact on Hong Kong educational development stated:

And they are [examinations] used to establish a ranking order among students as a basis for allocating a small number of places among a large number of applicants possessing the minimum qualification required. (Llewellyn Committee, 1982)

This situation has not changed and only the very best students who obtain high grades in public examinations go to university for further education that prepared them for entering professions (Lau, 2005). Though the competition for tertiary places has been lessened in recent years, the emphasis on examination for selection purpose is still much stronger than in some other places (Biggs, 1996). This indicates that the main practical purpose of public examinations in Hong Kong has been to select high-ability students for university admission (Chang, 2004). Thus, the wash-back effect on schools has been great. Schools have focused on drilling their students to prepare for the examinations, often at the expense of teaching and learning (Lau, 2005, p.195). Education in the classroom, therefore has been affected by public examination (Fullilove, 1992). Choi (1999, p.412) indicated that students sometimes stop their teacher from teaching topics not closely related to examination content. Chang (2004) suggested that a fair and reliable norm-referenced public examination was still a “must” in Hong Kong society. This is the background to efforts at reforming the system and the challenges faced by policy makers. It is always extremely difficult to change these well-established values, habits and modes of operation in our society.

The Hong Kong Examinations and Assessment Authority (HKEAA), however, made significant attempts to address these issues. Davison (2007) linked the introduction of SBA in the Hong Kong Certificate of Education (HKCEE) in the 2005/2006 academic year to Hong Kong’s broader curriculum reform agenda as outlined above. She located the initiative specifically as an attempt to take high stakes assessment (that counted for 15% of the final examination grade) and use it for both formative and summative purposes. This linking together of formative and summative assessments has been an important thread in the promotion of assessment for learning and supported by some very strong advocates of assessment for learning (Harlen & James, 1997; Harlen, 2005). In an important sense, the overall strategy of HKEAA in mainstreaming SBA as part of the new senior secondary curriculum was to label all assessment

as “assessment for learning”. This is a clever strategy but it has not been enough to declare SBA a success in Hong Kong.

I want to question, however, whether SBA can really be used as ‘assessment for learning’. My view is that high stakes SBA is , by definition, always linked to ‘assessment of learning’ since it contributes to a final assessment of student learning that is then used for decision making purposes. SBA has instrumental purposes rather than educational purposes - the results of examinations containing SBA components still serves the selection process in Hong Kong. What is more, as Lee (2007) has shown, teachers are not always able to distinguish between the formative and summative purposes of classroom assessment. Confusing the purposes of assessment will inevitably confuse students, teachers and other stakeholders and indeed has the potential to make SBA even more stressful when its purposes are uncertain or at best ambiguous.

Davison (2007) argued cogently that the results of the SBA oral assessment in English could be fed back to students to help them improve while at the same time being sent to the examinations authority to be included in the final examination score³. Yet Cheng, Andrews and Yu (2011) have shown that many students were unable to differentiate these purposes and for them SBA was simply like another examination. Davision (2007) herself shows that for many teachers there was a need to rehearse and practice students so they would do well on the SBA tasks – just as in an examination. Theoretically, Davison’s(2007) position makes good educational sense, but practically, and indeed I would argue culturally, it does not. When SBA is to serve instrumental purposes such as contribution to a final examination score, it seems to me better to treat it as such. I shall come back to this issue in the next part of the paper that will consider the cultural construction of assessment in Asian contexts.

An important aspect of Hong Kong’s SBA initiative was that it also represented a move from norm referenced to standards referenced assessment. Thus in the situation described by both Davison (2007) and Cheng et.al., (2011) students were being assessed against criteria and not against each other. While this started with the HKCEE examination as described above it was

³ School based results were actually moderated in relation to school performance in the public exam so the school based results may be changed depending on the moderation process.

also a principle designed to influence the new examinations system – a distinct break with the old British A-Levels examination system. Or (2008) studied the extent to which Hong Kong English teachers used the standards for HKCEE English and while she found considerable potential for such use teachers themselves did not seem to be able to integrate the standards into teaching and learning. This is understandable and as the standards are explained by HKEAA for the Hong Kong Diploma of Secondary Education (HKDSE) examination, they take on quite a technical tone despite the existence of qualitative descriptors. (HKEAA, 2013).

To summarise this section of the paper, what is clear from the Hong Kong case is that SBA was linked to a broader education reform agenda of which assessment in general and SBA in particular was a part. This led to SBA being conceptualised as “assessment for learning”, despite its high stakes nature and its relative newness as a feature of the Hong Kong education system. I have already expressed some concerns about this linking of SBA with assessment for learning. In the following section I shall explain why I have these reservations and why I think they are important.

Locating high stakes SBA in the social, cultural and educational contexts of Asian societies

In this section of the paper I want to make three key points:

1. Assessment policies are embedded in broader social contexts;
2. Teachers in Chinese societies have distinctive conceptions of assessment;
3. There are definitional issues relating to “assessment for learning” and “assessment of learning”

Assessment policies and their cultural contexts

This is by no means a new issue. My colleagues and I raised in several years ago (Kennedy, Chan, Fok & Yu 2008), it was reinforced in Kennedy & Lee (2010) and it has been nicely extended by Carless (2011). Examinations used for selection purposes are part of the history of China and belief in them as selection mechanisms has spread to most parts of the region. In this

sense examinations serve meritocratic purposes. Examinations are part of the social structure of many Asian societies selecting the best students to perpetuate the same structures and providing pathways to elite universities. The benefits of a university education are such that parents from many Asian countries will sacrifice a great deal to ensure that their children stay in the race and compete well. But isn't it the same in all countries that ration university places?

I would want to argue that in many Asian countries the pressures are different. Take, for example, the private tutoring industry in Asia. Bray & Kwok (2003, p. 614) have pointed out that “.. a major purpose of tutoring is to help pupils to gain qualifications, ... demand tends to increase close to the major public examinations, and then abruptly to decline once the examinations are over.” It has also been pointed out that from the point of view of students themselves, cram schools provided “shortcuts to learning, thorough past examination paper analysis, and even seemingly reliable open examination tips in Hong Kong, Taipei and Tokyo (Kwok 2004, p.70). Suffice it to say that the high stakes nature of examination in Asian countries leads to these kinds of social practices usually not seen elsewhere. This is because examinations are seen as distributors of merit and therefore very much prized and valued.

Yet private tutoring simply provides a means of doing well in examinations, its importance is more in the cultural purposes it serves especially in Confucian contexts Doing well in examinations relates in the first place to the way individuals see themselves . As I have written previously : (Kennedy, 2011, p 11):

Perhaps more important is the view that has been advanced by Li (2009, p.49) that for Chinese students “perfecting oneself morally and socially” is a fundamental purpose for learning. It is not the only purpose but it is ranked as the first purpose. This is consistent with Lee’s (1996) description of Confucian learning values in which self perfection plays a very important role. Thus not only does the immediate classroom context support Chinese learners but so too does a tradition that is

thousands of years old. Li (2009, p.61) talks about “learning virtues”: “resolve, diligence, endurance of hardship, perseverance and concentration”.

This motivation to work hard and do well is expected of Confucian families and students spending many hours after school with private tutors can be seen as a reflection of this particular value. As Sorenson (1994, p.14) put it almost two decades ago with specific reference to Korea, “young South Koreans’ educational success will be seen less as a matter of curriculum, class size, and educational technique than as a consequence of how education is embedded in the fabric of Korean society. It is this “embeddedness” of the importance of education in families, schools and individuals that makes the difference in Confucian societies. Doing well is not just a matter of succeeding at the “race” – it is a matter of honour and respect for the family. Getting into university is important: the honour that this brings to the family is more important. Thus success is a cultural imperative and private tutoring has become an important social institution that facilitates this imperative.

Chinese teachers have distinctive conceptions of assessment

A collaborative research project between researchers at the Hong Kong Institute of Education and South China Normal University has established that teachers in Hong Kong and Guangzhou have different conceptions of assessment from their peers in New Zealand and Australia and also show a number of cross border differences as well (Gao & Kennedy, 2011; Brown, Hui, Yu, & Kennedy, 2011; Brown, Kennedy, Fok, Chan, & Yu 2009). A key finding of these studies is the stress placed by teachers in both Hong Kong and Guangzhou on accountability as a motivation for assessment and improvement as a key outcomes for students. Student accountability is seen to be facilitated through examinations as is improvement. It is this link from accountability to improvement and examinations that characterizes Chinese teachers conceptions of assessment but is absent in conceptions of assessment held by teachers in Australia and New Zealand. This relationship is likely not just a reflection of educational

values – but of broader social values. As Brown, Kennedy, Fok, Chan & Yu (2009, p. 358) commented , “there is abundant evidence that high expectations for success and social improvement through examinations play a very significant role in the lives of Chinese families.” Such a view was supported earlier by Gow, Balla, Kember, & Hau (1996, p. 112) who commented in relation to Chinese students that “they work hard and generally attribute their academic performance more to their effort than to ability...They learn as much as they can in school so as to fulfill their duties towards their parents”.

The work to date on cultural conceptions of assessment, at least as far as this region is concerned, has been with teachers in Chinese contexts – Hong Kong, Guangzhou and other parts of Mainland China. The extent to which it is generalizable to other cultural contexts remains to be tested. Yet what is important to understand is that culture matters – Western imported educational innovations must always run up against cultural values and when we are searching for explanations about resistance to change we should not forget the significance of cultural values which will always trump new educational ideas and innovations.

Definitional issues – can we always recognize ‘assessment for learning’ and ‘assessment of learning’?

The Assessment Reform Group in what was billed as its “last publication” (TLRP, 2009) made very clear distinctions between the two forms of assessment: assessment for learning is “the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there”. Ten principles of assessment for learning have been developed and the gist of these is that assessment is very much a pedagogy or at least pedagogical in nature linked to supporting students in their learning journeys. This is opposed to different kinds of assessment for learning

– either internal or external summative assessment (Harlen, 2007). While the point is made that internal summative assessment can be used for formative purposes, this might not always be assessment for learning. This is not just definitional – assessment for learning is about more than providing feedback, it is about helping students get to where they want to go either with the support of their teachers, or under their own guidance or even the guidance of their peers, Summative forms of assessment, especially in the context of SBA where the marks count in an examination usually cannot meet the rigid criteria that defines true assessment for learning.

Summary

The point I want to make is a simple one: it is more than likely that assessment , examinations and social mobility are strongly linked in the minds and indeed the hearts of Asian students and their families. Changing the rules in the form of introducing SBA introduces not just a new educational idea; it raises all kinds of questions about reliability, validity, fairness and transparency. It leads to community questioning of the capacity of teachers to carry out such assessments and their honesty in doing so. It is for this reason in Hong Kong that the Independent Commission Against Corruption had to lay down some principles for the operation of SBA. For some reason, examinations are seen to be fair: SBA introduces an element of uncertainty and therefore a risk to the social outcomes of schooling which represent such significant parental aspirations. Teacher concerns about SBA might appear to be technical: in the end, however, it seems to me they are much more deep seated since SBA is often seen to challenge what society values most; an open, fair and transparent examination system.

Can SBA succeed in Asian contexts?

This is as much a rhetorical question as one to which there can be a neat and tidy answer. The Hong Kong experience is mixed – in the recently released *Progress Report on the New Academic Structure Review* (Curriculum Development Council et. al., 2013) there remains a commitment to SBA . Yet in light of student and teacher workload, and some evidence of drilling for assessments, SBA will be “cancelled in three subjects, deferred in 9 subjects , streamlined in a number of other subjects and replaced with a practical exam in another” (Curriculum Development Council, et. al., 2013, p. 46.). These changes were summed up under the general point that “SBA will be streamlined according to its necessity and suitability for

various subjects, with the pace of implementation adjusted to meet the needs of schools and their students (Curriculum Development Council, et.al., 2013, p.12).

What this suggests, in the Hong Kong context at least, is that SBA is difficult but not impossible. The extent to which it remains a feature of future end of school examinations will be a matter to watch. Yet one thing is clear: SBA no longer seems to be promoted as a form of assessment for learning s was being done when it was first introduced. This, I think is wise, because of the reasons outlined above. SBA is not assessment for learning in the way that has been outlined in this paper and by the Assessment Reform Group. Mistaking different purposes of assessment probably does not help the development and implementation of SBA.

At the same time, the Assessment Reform Group has been very clear about the standards that should be applied to SBA when it contributes to external assessment (TLRP, 2009, p. 15):

the credibility of the judgments made by teachers will need to be manifestly consistent and unbiased. Systems will be needed to ensure that all teachers engaged in making judgments in a particular context are working in comparable ways to an agreed set of criteria and standards.

For any context in which a much larger number of teachers are making judgments independently of each other, a more sophisticated infrastructure of guidance, training, support and cross-checking is required if the quality of those judgments is to be assured.

For all assessment, pupils, parents and teachers need assurance that the results for a particular pupil are comparable across different teachers in a school and between schools. Few schools are able, by themselves, to audit inter-school comparability, but they should have procedures in place to check intra-school comparability, to be followed up by inter-school moderation. In secondary schools, all subject departments should have a clear and documented assessment policy including specifications for the assessment instruments to be used, for ensuring validity, for resolving differences of opinion between teachers, and for procedures to be used to check the assignment of levels or grades.

These are very stringent requirements that should largely be the responsibility of education systems. But schools also need to be geared up to handle SBA in a way that may not apply to ordinary classroom assessments. The teacher education requirements are very clear and I wonder whether our teacher education programmes have incorporated these into preservice teacher education or whether sufficient emphasis is given to the inservice education of teachers who are responsible for SBA.

Conclusion

I want to end this paper on what has been a theme throughout: can SBA engender enough community support to be considered a fair and reliable component of the region's education systems? Examinations have a long history in the region and have developed community support as instruments that are seen to guarantee fairness, transparency and objectivity. The challenge for SBA is to be considered in the same way. Developing teacher capacity, building the kind of assessment systems referred to by the Assessment Reform Group and ensuring that assessment standards for SBA are as strong as they are for examinations are resource intensive and long term objectives. The validity offered by SBA must be well established and unable to be secured in other ways. SBA cannot be just faddish – the latest educational trend. It must recognize the cultural context in which it will operate and demonstrate how it meets the needs of that context. Winning the hearts and minds of students, teachers and the community is a big challenge – but it is one education systems in the region need to win if SBA is to remain a feature of the region's educational landscape.

References

- Assessment Reform Group. (1999). *Assessment for Learning: Beyond the Black Box*. University of Cambridge: School of Education.
- Biggs, J. (1996). The assessment scene in Hong Kong. In J. Biggs (Ed.), *Testing: to educate or to select: education in Hong Kong at the crossroads* (pp.3-12). Hong Kong: Hong Kong Educational Publishing Co.
- Bray, M & Kwok, P (2003). Demand for private supplementary tutoring: conceptual considerations, and socio-economic patterns in Hong Kong. *Economics of Education Review*, 22 (6), 611–20.
- Brown, G. T. L., Hui, S. K. F., Yu, W. M. & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50 (5-6), 307-320.

- Brown, G. T. L., Kennedy, K. J., Fok, P. K., Chan, J. K. S., & Yu, W. M. (2009). Assessment for improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy and Practice*, 16(3), 347-363.
- Broadfoot, P., & Black, P. (2004). Redefining assessment? The first ten years of *Assessment in Education*. *Assessment in Education*, 11(1), 7-27.
- Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. New York: Routledge.
- Chang, L. (2004, Nov 27). Structural reform overdue but wisdom of new tests is doubtful. *South China Morning Post*, p. Edu4.
- Cheng, J. (2004, Nov 18). Listen to the voice of teachers. *South China Morning Post*, p.17.
- Cheng, L., Andrews, S. & Yu, Y. (2011). Impact and consequences of school-based assessment(SBA): Students' and parents' views of SBA in Hong Kong. *Language Testing*, 28(2), 221–249.
- Cheung, D. (2001). School-based assessment in public examinations: identifying the concerns of teachers. *Education Journal*, 29(2), 105-123.
- Choi, C. C. (1999). Public examination in Hong Kong. *Assessment in Education*, 6(3), 405-417.
- Curriculum Development Council, Hong Kong Examination and Assessment Authority & the Education Bureau. (2013). Progress Report on the New Academic Structure Review - *The New Senior Secondary Learning Journey – Moving Forward to Excel*. Accessed from <http://334.edb.hkedcity.net/doc/eng/FullReport.pdf>
- Curriculum Development Council (2001). *Learning to learn: The way forward in curriculum*. Hong Kong: CDC
- Davison, C. (2007). Views from the chalkface: English language school-based assessment in Hong Kong. *Language Assessment Quarterly*, 4(1), 37–68.
- Education Commission. (2000). *Education blueprint for the 21st century: Learning for life, Learning through life - Reform proposals for education system in Hong Kong*. Hong Kong: Education Commission.
- Fullilove, J. (1992) The tail that wags. *Institute of Language in Education Journal*, 9, 131–47.
- Gao, L.B. & Kennedy, K. (2011). The Chinese mainland teacher evaluation Appearance: types, characteristics and enlightenment. *Education Journal*, 7 (1) , 39-47. (in Chinese)
- Hau, K. T. (2004, Dec 2). Better to change university entrance requirement than using school-based assessment (in Chinese). *Mingpao*.

Hong Kong Examinations and Assessment Authority [HKEAA]. (2005). School-based Assessment (SBA): a brief introduction and response to major concerns. Accessed from http://eant01.hkeaa.edu.hk/hkea/topper_tas.asp?p_coverdown=tas_1.html

IBM Corporation. (2003). *Strategic review of Hong Kong Examinations and Assessment Authority*. Hong Kong: Hong Kong Examinations and Assessment Authority.

Eisner, E. (1991). *Enlightened eye: Qualitative inquiry and the enhancement of educational practice*. New York: Macmillan.

Gow, L., Balla, J., Kember, D & Hau, K. T. (1996). The learning approaches of Chinese people: A function of socialization processes and the context of learning? In M. Bond (Ed.), *The Handbook of Chinese Psychology* (pp. 109-123). Hong Kong: Oxford University Press.

Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *The Curriculum Journal*, 16 (2), 207 – 223.

Harlen, W. (2007). *Assessment of learning*. London: SAGE.

Harlen, W. & James, M. J. (1997). Assessment and learning: differences and relationship between formative and summative assessment. *Assessment in Education*, 4(3), 365-380.

HKEAA. (2013). Standards referenced reporting. Accessed from <http://www.hkeaa.edu.hk/en/hkdse/srr/>

Kennedy, K. (2011). Teacher quality and its cultural contexts-What can the West learn from the East? *Journal of Research, Policy and Practice of Teachers and Teacher Education*, 1 (1), 8

Kennedy, K., Chan, K. S. J., Fok, P. K., & Yu, W. M. (2008). Forms of assessment and their potential for enhancing learning: Conceptual and cultural issues. *Educational Research for Policy and Practice*, 7 (3), 197-207.

Kennedy, K., & Lee, J.K.C. (2010). *The Changing Role of Schools in Asian Societies – Schools for the Knowledge Society* [Paperback Edition]. New York: Routledge.

Kwok, P. (2004). Examination-oriented knowledge and value transformation in East Asian cram schools', *Asia Pacific Education Review*, 5 (1), 64-75.

Lau, C. K. (2005). Between egalitarianism and elitism: media perception of education reform in post-1997 Hong Kong. In L. S. Ho, P. Morris, & Y. P. Chung (Eds.), *Education reform and the quest for excellence: the Hong Kong story* (pp.191-216). Hong Kong: Hong Kong University Press.

Lee, I. (2007). Feedback in Hong Kong secondary writing classrooms: Assessment for learning or assessment of learning? *Assessing Writing*, 12, 180–198.

Llewellyn Committee (1982). *A Perspective on education in Hong Kong: Report by a visiting panel*. Hong Kong: Government Printer.

OECD. (2005). *Formative Assessment: Improving Learning in Secondary Classrooms*. Paris: OECD.

Or, J.Y.K. (2008). Formative use of standards-referenced assessment in Hong Kong. Unpublished MEd thesis, Faculty of Education, University of Hong Kong. Accessed from <http://hub.hku.hk/handle/10722/56127>

Sorensen, C. (1994). Success and education in South Korea. *Comparative Education Review*, 38 (1), 10-35 .

Teaching and Learning Research Programme (TLRP). (2009). *Assessment in schools - Fit for purpose?* A Commentary by the Teaching and Learning Research Programme. Accessed from <http://www.tlrp.org/pub/documents/assessment.pdf>

Yip, D. Y., & Cheung, D. (2005). Teachers' concerns on school-based assessment of practical work. *Journal of Biological Education*, 39(4), 156-162.

Yung, B. H. W. (2001). Three views of fairness in a school-based assessment scheme of practical work in biology. *International Journal of Science Education*, 23(10), 985-1005.

Yu, F.W.M., Kennedy, K., Fok, P.K & Chan, J.K.S. (2009). Assessment reform in basic education in Hong Kong: The emergence of assessment for learning. Paper presented at the 32nd Annual Conference of the International Association for Educational Assessment, Singapore. Accessed from: http://www.ied.edu.hk/fpece_project/QEF/Download%20area/4_Singapore.Flora.paper.pdf

School-Based Assessment, Standards, Teacher Judgement and Moderation: Messages from Research

Professor Val Klenowski
Queensland University of Technology
Val.Klenowski@qut.edu.au

Introduction

In Australia, states and territories are implementing a National Curriculum and Achievement Standards, which present many challenges for teachers' work. Australia is not the only country that has introduced the use of standards in recent years. Countries such as Canada and New Zealand have also introduced standards with more teacher assessment. In this paper I will focus on the implications of standards-based assessment reform for students, teachers and systems. In particular, I will highlight the significance of teacher judgement and moderation in standards-referenced assessment for accountability and improvement of teaching and learning. To conclude some suggestions for practice and policy will be summarized with the important message for policy to provide support for teachers in times of major educational reform.

Australian Context

In Australia there are six states and two territories and each jurisdiction is responsible for curriculum development, implementation, assessment and reporting. At the federal level, the Australian Curriculum, Assessment and Reporting Authority (ACARA) is responsible for the continuing development of a national curriculum made up of content descriptions and achievement standards.

There is a range of assessment types that are practised in Australia. Students receive an A to E report card every semester in every year in every subject. School-based testing and assessment comprises the National Assessment Program – Literacy and Numeracy (NAPLAN) tests conducted in May, students receive their results in September, and the school receives diagnostic information in December or January. In Queensland up until 2013, the Queensland Studies Authority (QSA) developed the Queensland Comparable Assessment Tasks or QCATs that have been conducted in Years 4, 6 and 9 (<http://www.qsa.qld.edu.au/3163.html>). These tasks were designed to assist teachers to understand the qualities needed in student work indicative of the national achievement standards. The tasks are more like rich tasks and are intended to demonstrate to teachers how tasks can be designed to provide the opportunity for students to demonstrate understanding as well as skills such as critical thinking, reflection and investigation. Other localised forms of testing include classroom tests, diagnostic tests using Progressive Achievement Tests in

Reading (PAT-r) developed by ACER. Parent assessment also takes place through observation and after school study.

Queensland

The Queensland Studies Authority recognised the importance of supporting teachers when standards were first introduced. A Curriculum, Assessment and Reporting Framework was established in this state at that time when the Australian Government in 2006 moved to implement an A to E reporting system for the purpose of providing standards-related information to parents. This initiative focused on improvement in teaching and learning by providing clarity about the Key Learning Area syllabus documents and the consistency of what was taught across the state. This initiative was also taken to prepare for the implementation of a standards-referenced system. The framework was intended to support teachers by providing defined standards and guidelines to:

- help build their capacity in assessment
- make judgements about student work
- use information to provide formative feedback and
- to maintain comparability of reported student achievement to parents and carers.

Achievement Standards

The Australian Curriculum has been described as futures-oriented in that it is designed for a complex society with workers competing in a global market, needing to know how to learn, adapt, create, communicate, interpret and use information critically. The Achievement Standards comprise a written descriptor plus annotated student work samples, to indicate an expectation of the quality of learning that students should typically demonstrate by a particular juncture in their schooling. The provision of standards and annotated samples of work has been found to be necessary but insufficient for teachers to develop consistency in their judgements of student work and comparability in their judgements when using the standards. This is because the standards are socially constructed and historic in nature. Given these characteristics standards can achieve acceptance at a point in time, which can change over time.

The way standards are represented conveys expectations of quality and levels of performance. Standards provide a common set of stated reference points, as in the Australian Achievement Standards. They are published features of quality against which teacher judgement can be held accountable or scrutinized. In the teaching and learning cycle teachers are encouraged to share the standards with students to provide information about the expected qualities they are

aiming for. In this way the standards are linked to teacher feedback and student self and peer assessment. However, it is important to emphasise that the provision of standards and annotated work samples are necessary but insufficient.

Here is an example of the Year 9 (14 year old) Achievement Standard for Science. This standard is designed to indicate the quality of learning that students should typically demonstrate by the end of Year 9. The Achievement Standard comprises the written description.

By the end of Year 9, students explain chemical processes and natural radioactivity in terms of atoms and energy transfers and describe examples of important chemical reactions. They describe models of energy transfer and apply these to explain phenomena. They explain global features and events in terms of geological processes and timescales. They analyse how biological systems function and respond to external changes with reference to interdependencies, energy transfers and flows of matter. They describe social and technological factors that have influenced scientific developments and predict how future applications of science and technology may affect people's lives.

Students design questions that can be investigated using a range of inquiry skills. They design methods that include the control and accurate measurement of variables and systematic collection of data and describe how they considered ethics and safety. They analyse trends in data, identify relationships between variables and reveal inconsistencies in results. They analyse their methods and the quality of their data, and explain specific actions to improve the quality of their evidence. They evaluate others' methods and explanations from a scientific perspective and use appropriate language and representations when communicating their findings and ideas to specific audiences.

Such standards are important for informing teaching and learning in terms of the development of assessment tasks. Teachers are provided with student work samples that are intended to communicate the expectations of quality described in the standard and which reflect the design of the assessment task. Each work sample includes the task, the student's response and annotations identifying the quality of learning evident in the student's response in relation to relevant expectations of the standard.

The Achievement Standards in Australia now need to be validated over a period of least two years of teacher use of these standards. Teachers need to be observed using the standards to decide whether they are making consistent judgments using the achievement standards and to examine how teachers arrive at consistent judgements. Currently the standards are not refined to identify the correct pitch, regarding the level for use by teachers with their students' work. In the advice from ACARA it is stated that together the Achievement Standard plus the annotated work samples help teachers to make judgements about whether students have achieved the standard. However, an essential process is missing from this guidance.

Representation of Standards

The way in which standards are formulated influences not only their representation but also suggests a particular approach to judgement. In developing achievement standards it is important to use the variables of concept/skill, verb/cognition and degree/qualities. Too often only two of these variables are used such that concept/skill and verb/cognition are incorporated without a degree/qualifier (the “how well”). Too often the standard describes *more* and *different* work for each level, which means that what is being assessed at a particular level of the standard is different from what is being assessed in the other levels of the standard. As a result rather than providing a basis for differentiating how well the same concept/skill has been demonstrated A-E, each of the standards will be about a different concept/skill and verb/cognition. For example, if to show the differences in the levels of a standard words such as ‘evaluates’ or ‘analyses’ are used to describe an A level response and ‘identifies’ is used for a D level response this will indicate that what is being assessed at the A level is different from what is being assessed in the other levels of the standard. What occurs in this context is rather than providing a basis for differentiating how well the same concept/skill has been demonstrated (A-E), each of these standards is about a different concept/ skill and verb/cognition. This makes it very difficult to set assessment tasks. This is because a task requiring evaluation is very different from one that requires explanation (QSA, 2012).

The representation of the standards as in a matrix format or as continua will also have implications for how teachers approach their judgement practice. For example, research found that when a matrix format was used teachers were more likely to adopt an analytic approach to judgement, however when the standards were represented as continua the approach suggested was more holistic (Klenowski & Wyatt-Smith, 2010a).

Implications for practice

The implications for teachers of the implementation of a standards-referenced system and a futures-oriented curriculum is that teachers can no longer rely solely on assessment formats such as paper and pencil tests as these do not allow students to demonstrate fully the extent of their knowledge, the depth of their understanding and the sophistication of their skills.

In England an expert enquiry into the Key Stage 2 (students are aged 11) Science tests recommended that they be abandoned. This decision relates to Professor Philip Adey's research findings that:

The general cognitive foundation of 11 and 12-year-olds [had] taken a big dip. There [was] a continuous decline for the last 30 years and ... [was still continuing in 2006]
(Griffiths, 2006)

Possible explanations for why this occurred are numerous but one that is pertinent to the importance of the shift to assessing skills and the use of teacher judgement using standards is that:

By stressing the basics - reading and writing - and testing like crazy you reduce the level of cognitive stimulation. Children have the facts but they are not thinking very well. ... And they are not getting hands-on physical experience of the way materials behave. (ibid)

In this context 'teaching to the test' resulted in a reduction in hands on practical tasks, which in turn *led to students' conceptual skills actually decreasing over the next ten years*. So while the test results improved year on year, *the learning and understanding decreased*.

Teacher Judgement

Teachers in Australia with the use of Achievement Standards and the national curriculum now need to assess students' *application of their understanding and the development of their skills such as inquiry, analysis, investigation and reflection*. Such reform has implications for how teachers make their judgements of student work. They need to draw on the following intellectual and experiential resources when making those judgements. Sadler (1998: 80-2):

- Superior knowledge about the content or substance of what is to be learned
- Sound knowledge of criteria and standards (or performance expectations) appropriate to the assessment task
- Evaluative skill or expertise in having made judgements about students' efforts on similar tasks in the past
- A set of attitudes or dispositions towards teaching, as an activity, and towards learners, including their own ability to empathise with students who are learning,

their desire to help students develop, improve and do better, their personal concern for the feedback and veracity of their own judgements, and their patterns in offering help.

Moderation

Using standards for the first time can be challenging for teachers as confidence builds over time through their use and standards acquire meaning through use over time. This is because standards when written as verbal descriptors require interpretation and application in a community of practice. That is, moderation practice. The introduction of standards and more performance oriented tasks requires teachers to engage in moderation because it supports both consistency and comparability in teacher judgements.

Moderation practice is important for the development of shared understandings about the standards through application and use. Teachers also deprivatise their practice in moderation meetings as they have to defend their judgements and provide the evidence for arriving at a particular judgement and/or awarding student work a particular standard or grade level. It is therefore important that in a standards-referenced system teachers have the opportunity to meet for the purposes of learning and teaching improvement and accountability in terms of quality assurance and comparability.

States and territories have responsibility for assessment and reporting and in working with ACARA aim for a more nationally consistent approach. Each jurisdiction has developed its own plan for implementation. For example, the Queensland Studies Authority has developed curriculum and assessment products to support schools to assess student work and to report twice yearly using an A-E scale using the Australian curriculum. The Learning Area Standards Descriptors (LASDs) elaborate on the Australian Curriculum achievement standards. They use the two dimensions of the Australian Curriculum achievement standards of understanding and skills. They also describe the qualities in student work using a five-point scale (A-E) and are designed to focus teacher attention on the valued features drawn from the Australian Curriculum achievement standards. These descriptors will now be validated to investigate how well the LASDs align to the Australian Curriculum and how well they assist teachers in the following three key areas:

- Making an on-balance judgement based on a folio of student work at the end of a reporting period
- Developing task-specific guides to making judgements
- Developing a balanced assessment program and individual assessment tasks.

Research Findings

Standards, moderation and teacher judgement are all required. This involves teachers identifying evidence in student work, use of exemplars to guide judgement practice for consistency and comparability and moderation to reach agreement and identify implications for learning improvement. There are four conditions that have been identified for the use of standards to serve to improve learning and to fulfill an accountability function in system reporting of valid, reliable and equitable practice.

These conditions are:

- Clarity around the purposes and functions
- Understanding the representation of the standards
- Moderation for quality assurance, comparability and learning improvement
- Assessment community and assessment culture (Klenowski & Wyatt-Smith, 2012b).

Standards need to be validated through interpretation and negotiation in moderation practice and should be empirically derived. Moderation, a social practice of exchanging views of quality for the purpose of comparable and consistent judgement, is essential in a standards-referenced system to achieve inter-rater reliability. Moderation is particularly important when an on-balance judgement is made of folios of work and to address the threats to validity through ‘construct-irrelevant variance’ or ‘construct under-representation.’

Conclusion

In Queensland the Queensland Studies Authority has been instrumental in providing the necessary framework which comprises guidelines and advice, resource development, professional development that are communicated to the teachers, schools, stakeholders using ICT infrastructure. In anticipating the teachers' needs at the time of major curriculum and Achievement Standards the QSA has worked with researchers to develop the necessary policies, practices and resources to support teachers. One such resource has been the development of assessment and reporting advice, guidelines and principles to guide the effective curriculum and assessment planning. These principles include:

- high expectations for all students
- alignment of teaching and learning, and assessment and reporting
- standards- and school-based assessment for learning
- balance of informed prescription and teacher **professional judgement**
- planning a balanced assessment program for developing the components of an assessment folio for the purpose of making an on-balance judgement

- developing assessment techniques, formats and conditions appropriate for quality assessment for each learning area
- monitoring student achievement and providing feedback.

Teachers in the education system are viewed as the primary change agents, who through judgement practices that are integral to the requirements of assessment tasks and expectations of quality performance, are best placed to identify important steps for students to improve in their learning and to develop useful insights about how best to change pedagogy to meet students' particular learning needs. The QSA supported research and development to learn of the level of support required for teachers in the implementation of a standards-referenced system. However, teacher judgement remains under-researched and remains in its infancy

References

- Griffiths, S. (2006) Failing to teach them how to handle real life, Times Online, Accessed 13 November, 2009 from <http://www.timesonline.co.uk/tol/news/article721863.ece>
- Klenowski, V. & Wyatt-Smith, C. (2010a) Standards-driven reform years 1-10 : moderation an optional extra? *Australian Educational Researcher*, 37(2), pp. 21-40.
- Klenowski, V. & Wyatt-Smith, C. (2010b) Standards, teacher judgement and moderation in contexts of national curriculum and assessment reform. *Assessment Matters*, 2, pp. 107-131.
- Queensland Studies Authority (2012) Meeting Paper, Assessment, Moderation and Certification Committee meeting (Friday 31, August 2012), Brisbane: Queensland Studies Authority.
- Sadler, D.R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy and Practice* 5, 1: 77–85.

SCHOOL-BASED ASSESSMENT : TRANSFORMATION IN EDUCATIONAL ASSESSMENT IN MALAYSIA

HJH NORZILA MOHD YUSOF
EXAMINATIONS SYNDICATE
MINISTRY OF EDUCATION MALAYSIA

Introduction

Malaysia has introduced National Educational Assessment System (NEAS) since 2011 as part of its national educational transformation plans to achieve the aspiration of National Philosophy of Education in developing learners' physical, emotional, spiritual and intellectual abilities comprehensively and holistically. NEAS enhances meaningful learning through assessments of learners' profile, achievement, development and involvement through five various modes; School-Based Assessment (comprises of four components) and the Centralised / Public Examination. It is a move towards reducing the over reliance and over dependence on public examinations for certification.

Concept of School-Based Assessment in Malaysian Context

The concept of School-Based Assessment or more commonly known as Pentaksiran Berasaskan Sekolah (PBS) beginning Year 1 in 2011 and Form 1 in 2012 is a reengineering process of the educational assessment in accordance to the National Key Result Area (NKRA) agenda. The implementation of PBS paves the way to a meaningful assessment. It is characterised by its authenticity and robustness besides being holistic, integrated, low stake and comprising of quality assurance.

PBS consists of four major components, the School Assessment, the Central Assessment, the Physical Activities, Sports and Co-curricular Assessment and the Psychometric Assessment. Figure 1 shows the academic and non-academic components of PBS.

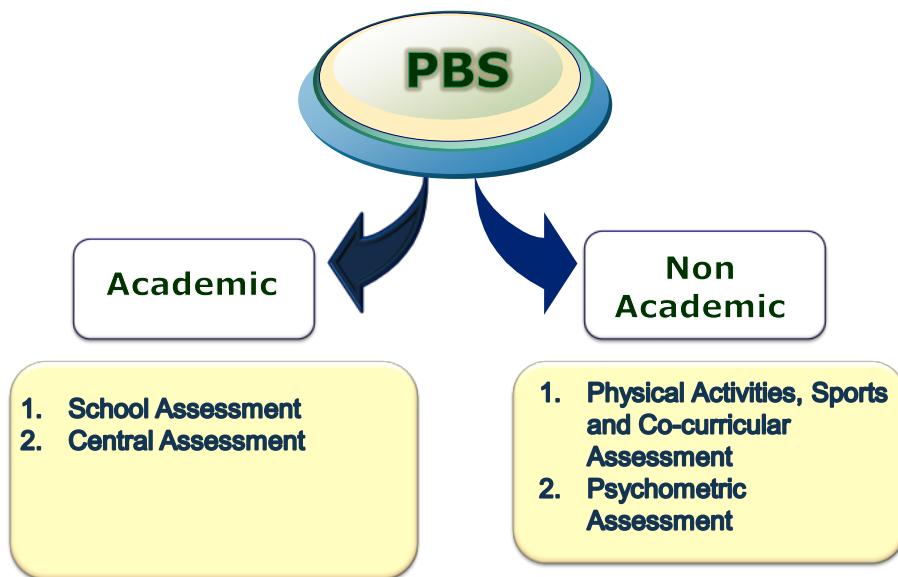


Figure 1 : The academic and non-academic components of PBS

School Assessment

School Assessment emphasises on collecting first-hand information about learners' learning based on curriculum standard, whereby teachers plan their assessment, prepare the instruments, administer the assessment, examine learners' responses and report their progress. Teachers could conduct formative assessment and provide constructive feedback and feed forward to learners. Teachers could also conduct summative assessments. Teachers are encouraged to develop various forms of assessments tasks that are authentic, contextualised and could enhance learners' higher order thinking skills and the 21st century skills such as creative, innovative, problem solving and decision making apart from instil moral values. Standard-referenced assessment is the basis of School Assessment to ensure that learners' performances are comparable to accepted world standards in various areas of knowledge, skills and competence. The Examinations Syndicate developed the performance standards based on the standard-based curriculum and they will be the guiding factor that helps teachers, learners and parents monitor learners' progress in learning. Statement of standard for each subject from year 1 to Year 6 is based on the objectives of the respective subjects. It is a generic explanation about the level of learners' learning development and growth.

The performance standards consist of descriptors and evidence for six bands or levels that learners have to achieve. The lowest is Band 1 and the highest is Band 6. Learners' performance is no longer marked by using marks or grades to indicate to the level of acquisition of knowledge and skills during teaching and learning. The descriptor refers to what learners have to know and are able to do as prescribed in the curriculum standards of respective subjects. Therefore, the descriptor for each statement of standard is different from year to year. The evidence refers to how learners demonstrate what they have acquired and what they are able to do.

Table 1 shows the framework of the performance standard while Table 2 shows the framework of the performance standards document used in the School Assessment.

BAND	STANDARD
1	Know
2	Know and understand
3	Know, understand and able to do
4	Know, understand and able to do in a systematic manner
5	Know, understand and able to do in an admirable manner
6	Know, understand and able to do in an exemplary manner

Table 1 : Framework of the Performance Standards

BAND	STATEMENT OF STANDARD	DESCRIPTOR	EVIDENCE
1	Generic explanation about the level of learners' learning development and growth (based on objectives of the subjects)	What learners have to know and are able to do (based on the curriculum standards of a particular subject for a particular year)	How learners demonstrate what they have acquired and what they are able to do
2			
3			
4			
5			
6			

Table 2 : Framework of the Performance Standards Document

Central Assessment

Central Assessment is another academic component in PBS which is conducted and administered by teachers in schools using instruments, scoring rubrics, guidelines, time line and procedures prepared by the Examinations Syndicate. Authentic assessment with various instruments and tasks is designed for Central Assessment to gather sufficient and more accurate information about learners' ability and performance. The process of administering this assessment will be monitored by the Examinations Syndicate and the PBS committee at school, district and state levels to ascertain that the scores awarded by the teachers comply with the rules, guidelines and procedures. Teachers have to submit learners' scores in this assessment to the Examinations Syndicate.

Physical Activities, Sports and Co-curricular Assessment (PAJSK)

Physical Activities, Sports and Co-curricular Assessment or PAJSK refers to the type of assessment which records learners' physical endurance and body mass index, and learners' participation, involvement and contributions in sports, co-curriculum and extracurricular activities. Such records serve as added value to learners' well-being and complement the introduction of the concept and policy of One Student One Sport or known as 1M1S (Sports Division, 2011). The information collected is to encourage learners to participate in physical and outdoor activities alongside with the academic achievement. This aspiration is very much desired in a multi-racial country like Malaysia in order to maintain a united, peaceful and harmonious nation. Furthermore, the record and report of these activities could also be used as added value to learners' application for further education and training. This gathering of information is an acknowledgement and recognition of learners' participation in sports and outdoor activities as part of learning process which enriches learners' knowledge, soft skills, and experience. Such activities are the driving factors to building a strong and rounded character in enhancing human capital as advocated in the National Philosophy of Education.

Psychometric Assessment

The fourth component of PBS is the Psychometric Assessment; a profiling assessment which emphasises on learners' learning inclination, attitudes, aptitude, interest and personality. This assessment enables teachers to identify learners' interest, innate ability and learning styles, and subsequently help them progress in their learning accordingly. The Examinations Syndicate prepare the psychometric test instruments and guidelines with the help of professional psychologists and counsellors. The Psychometric Assessment is conducted whenever needed in the upper primary level. School counsellors are responsible to administer the Psychometric Assessment in school and provide appropriate professional advice based on the results. However, as the results of personality inventory are confidential, only the results of the aptitude test are shared with the class or subject teachers to help learners in their learning process. As the assessment is used to help learners, the results of the assessment will not be used to influence learners' overall achievements or grades.

Quality Assurance

There is no doubt that quality assurance plays an important role in the success of any school-based assessment and so does it for PBS. Therefore, the Examinations Syndicate has set up a mechanism of quality assurance to ensure that the implementation of PBS is in accordance to its intended goals. The proposed quality assurance mechanism includes the mentoring, monitoring, moderating, and trekking measures to ensure the validity and reliability of the assessments. Schools will be monitored to ensure that the implementation of PBS is in accordance to procedures. Learners' performance will be moderated to ensure that there is uniformity and reliability in teacher's recording and reporting of an assessment. Teachers will be guided, facilitated and mentored in the process of conducting an assessment in school. To ensure the success of PBS in schools, study, feedback and reviews will be carried out as trekking tools to enhance the instrument and method of assessing learners through PBS. The quality assurance is the responsibility of the Examinations Syndicate and the PBS Committee at the state, district and school levels. In addition, the PBS Committee at various levels have to conduct PBS briefings,

meetings and trainings for those under their supervision, and prepare reports on the implementation of PBS to higher authority.

PBS Management System (*Sistem Pengurusan PBS – SPPBS*)

Judging from the four components of PBS, one would raise the question of the workload and clerical work that teachers have to endure in the implementation of PBS. In actual fact, teachers have been assessing learners since long time ago either formally through quizzes and tests or informally through question and answer in the classroom. PBS is introduced to allow all of these assessments that teachers have been doing to be recorded and reported in a more systematic way so that learners' progress and growth in learning could be monitored carefully. Subsequently, steps can be taken to improve learners' performance and acquisition of knowledge or skill. The question of workload and clerical work the teachers and public are worried about is addressed with the introduction of two computer applications which are the PBS Management System (*Sistem Pengurusan Pentaksiran Berasaskan Sekolah – SPPBS*) and the PAJSK Application. The SPPBS is meant for recording and reporting the PBS academic component, while the latter is for the non-academic component. These web-based systems were developed to help teachers record information about learners' acquisition of knowledge and skills in PBS. These systems can generate descriptive reports that can be given to parents, teachers or other stakeholders whenever required. This will reduce teachers' workload to record learners' learning manually and thus, lessen their clerical work.

Centralised Examination

The centralised examination for six main subjects (Malay Language, English Language, Science, Mathematics, Chinese Language and Tamil Language) will still be held at the end of Year 6 with a small percentage taken into account for the final score of learners' overall performance in primary schooling. However, this standardised paper and pencil exam will no longer be the only determining qualifier of learners' knowledge, skills and ability.

Issues and Challenges

After more than two years of implementation of PBS in primary schools and more than a year in the secondary schools, related issues and challenges have been identified through series of monitoring activities at various levels, questionnaires, dialogues and small studies conducted by the Examinations Syndicate. Among the issues are those related to man, machine, materials, methods and environment.

The main issue identified which is related to the man is school readiness in implementing PBS. It was found that school administrators and teachers still have difficulty in accepting the changes made in the policy. The reason being they do not get sufficient information on PBS and lack of relevant trainings. Another issue is teachers' skills which are found inadequate especially in the aspect of developing various assessment instruments other than written tests which they are used to. Materials on PBS are found to be insufficient for the teachers to refer to whenever they encounter a problem or have uncertainties to implement PBS.

Training all the teachers is a big challenge as it involves a large number, up to a few hundred thousands of them. The cascading training model applied in the initial trainings was found to be not effective enough as dilution of information occurred during the trainings. Subsequently, many teachers have different interpretations on how to carry out PBS in schools. The Examinations Syndicate then used the in-situ and hands-on approach method starting October 2012. In this training approach, the Examinations Syndicate officers will directly train the State and District Education officers together with the Head Teachers and the teachers of the identified district exemplary schools who will then train the Head Teachers and the teachers of the schools within the district. They are provided with PBS hands-on training modules and supporting materials on PBS as guidance and reference. The exemplary schools, identified master trainers in schools and the supporting material provided could be referred to as PBS implementation model. Coaching and monitoring after the training will be carried out to guide teachers while teachers' reflection on their PBS practice and feedback provided will enhance their understanding and confidence in implementing PBS.

The SPPBS, which is supposed to be helpful to teachers in reducing their workload has created problems to teachers instead and thus raised the machine-related issue. This is because teachers have difficult time accessing the system due to slow connections and bottled up server and hosting. Various Information and Communication Technology (ICT) divisions in the Ministry of Education have played their roles in rectifying the problems by upgrading the internet connections and providing a separate hosting for SPPBS. The problems have been resolved to a certain extent and teachers have already an easier accessibility to the SPPBS starting March 2013.

The environment-related issue in the implementation of PBS is the class size where the number of learners in each class could reach up to 50 in areas with high density of population such as the urban areas. Teachers in such schools will have hard time organising and managing learner-centred activites and assessments that cater for individual differences. On top of that they have to record learners' progress and development regularly. Physical development in the form of building more classrooms in a limited space of the school compound is sometimes not possible. Therefore, teachers are encouraged to apply collaborative teaching in the classroom so that PBS could still be carried out without burdening them too much.

The Examinations Syndicate has also done the outreach programmes to parents, inviting them to put PBS in their hearts and hence supporting the implementation of PBS. Emphasised on the descriptive reporting was one of the main agenda in the program so that parents know the meaning of scores that their child gets. Booklets on Frequently Asked Questions on PBS have been published and distributed to ensure uncertainties been answered.

Conclusion

It is hoped that the introduction of PBS in NEAS will be the turning point towards a progressive teaching and learning experience. The Ministry of Education expects the outcomes of PBS will inherently contribute towards achieving the objective of the National Philosophy of Education in developing human capital and could fulfil the aspiration to become a developed nation by 2020. PBS is hoped to provide more meaningful learning

and assessment which could equip learners with relevant and necessary knowledge, skills and attitudes to survive in this ever challenging world. It is hoped that by setting high standards, establishing measurable goals and monitoring learners' progress closely, learners' outcomes in education can be improved. PBS is a learner-centred approach towards learning and teaching. PBS should be able to make teaching and learning more fun and meaningful to learners.

This journey of PBS as transformation in educational assessment is going to be difficult, is going to be long, but if Ministry of Education don't bring about the change now, Malaysia may be left far behind. The Examinations Syndicate and the Ministry of Education have the biggest role to play in the success of PBS alongside the parents, stakeholders and everyone involved in the best interest out of learners' education in Malaysia.