課程與教學季刊　2010, 13(1), 頁 67～94

# 香港基本能力評估

莫慕貞

基本能力評估是香港特區政府面對 21 世紀全球化、知識爲本的社會和經濟等變化與挑戰的一個重要因應策略。本文旨在介紹香港基本能力評估（BCA）的設計原則和實施情況。「學會學習的評估」是基本能力評估設計的核心思想。基本能力評估包含基本能力評估之學生評估，和全港性系統評估兩個組成部分。前者是爲在第一至第三學習階段的中、英、數三科所提供的低風險、標準參照、自發性的校內在線評估。目的是促進校內教學效能。後者是一個由政府中央管理、全港學校參與的保密測試。目的是爲政府提供本港學生在這三科達到基本水平的數據，以作爲教育決策的準則。本文簡介基本能力評估的設計、評改方法、報告方式、標準制定、和公眾的接受程度等議題。

關鍵字：基本能力評估、全港性系統評估、香港、學會學習的評估、評估改革

作者現職：香港教育學院講座教授

# 1. Background of Assessment Reform in Hong Kong

The context of education in the 21st century is a complex one, characterised by globalisation, knowledge-based society and economy, strong expressed needs for education and changes in conceptions of learning. Since the turn of the century, many countries in the Asia Pacific region have initiated education reforms in response to the new developments in the conception of learning and in technology in the 21st century (Cheng & Townsend, 2000). Lifelong learning is recognized by major systems around the world as the necessary pathway for the future. For instance, the G8 governments and members of the European Commission declared their commitment to lifelong learning at the 2000 meeting. The Japan education aims to "enhance children's ability to think and learn for themselves"; Korea education aims to "raise a self-reliant individual equipped with a distinct sense of independence, a creative individual with a sense of originality"; Singapore government aims to education young people "to be innovative – have a spirit of continual improvement, a lifelong habit of learning and an enterprising spirit in undertakings"; Thailand education aims to "develop student's learning capabilities in the areas of self-learning, creative thinking and basic academic learning", and Hong Kong education aims to "enable every person to attain all-round development… so that s/he is capable of lifelong learning…and contribute to the future well-being of the nation and the world at large." The assessment reform in Hong Kong took place against this international context.

In the Asia Pacific Region, assessment reform is often seen by policymakers as leverage for educational change. Mok and associates (Mok, Gurr, Izawa, Knipprath, Lee, Mel et al., 2003) surveyed eight locations in the Region, including Hong Kong, Japan, Korea, Singapore and Taiwan, and found that reforming the assessment systems is one of the seven common trends that emerged in the education reform of these systems. Many of these countries had in the past been dominated by strong examination cultures which emphasise assessment for selection. Assessments in these countries tended to be high stake as assessment often took the form of a one-off examination and the outcome of which determined future prospects of education and employment of the examinees. Biggs (1996) observed that students in Hong Kong "focus their learning on what they think they will be tested on: the test becomes the curriculum" (p. 8). The negative backwash effect of high stake examinations on student learning have been strongly criticised by Hong Kong educators who advocated for change (e.g. Cheng & Townsend, 2000; Morris, 1985;

Pong & Chow, 2002; Yeung, 1996). An urge to reform the traditional assessment system has come also from the Education Commission in its EC Report Number 4, which states inter alia, "… if we are to develop an education system which provides for the different needs of students, we must be able to assess their individual strengths and weaknesses. We therefore firmly believe that the time has come for the development of an assessment system that would serve a formative function and which would enable the performance of students to be measured against agreed targets. (Education Commission, 1990, Section 5.2.3)". This Report gives impetus to the assessment reform in Hong Kong.

Indeed, there is broad consensus amongst education leaders in the Region that the time for change has arrived, and that change involves a paradigm shift from assessment of learning to assessment for learning. In 1998, the UNESCO published the Delor's Report, a report that has significant impact on reforms in the Region. The Report (Delors, 1996) highlights the changing global context of education and identifies education to be a total experience organised around four pillars of learning, namely, learning to learn, learning to live together, learning to do and learning to be, which together enable learning throughout life. Research in the USA (e.g. Autor, Levy, & Murnane, 2003) also found that new skills are required of graduates of the 21st century. Analysis on skill contents required of jobs as result of recent technological change in the USA between 1960 and 1998 found substantial increase in demands for expert thinking and complex communication together with drastic decrease in demands for routine work. Education has to change to meet new demands of the workplace.

In parallel with the changing conceptions of education are new research findings of factors contributing to learning. Of particular note are findings of the major review undertaken by Black and Wiliam (1998) which show that quality feedback is the crucial factor contributing to effective learning. If students are provided with quality feedback, they are more likely to learn better. Quality feedback has to be timely; provides a clear expectation of what the learner can do at each level; indicates what the learner is able to do; shows progress and identifies areas for improvement (Rowe, 2005). That feedback is crucial for effective learning is supported by parallel research in the USA. Major review of relevant literature by the American Psychological Association Board of Educational Affairs (1997) classified factors affecting learning into four categories, and first and foremost of these are cognitive and metacognitive factors. Formative assessment generates accurate and timely feedback that informs the learner about his/her own progress

and contributes to the learner's metacognition. The notion that assessment can be used formatively to generate diagnostic information for use by learners and their teachers to decide how much has been learned, what else needs to be learned, and how best to approach further learning is quickly embraced by educators and policymakers in the Region. Since the turn of the century, assessment for learning has became a powerful movement within the Region that sets out to change the traditional high-stake assessment of learning to low-stake assessment that yields evidence to inform students of progress and directions for subsequent learning. In the spirit advocated by Black and Wiliam (1998), the reform in the Region is to turn assessment from its curriculum domineer role (Biggs, 1996) into a servant of learning. Assessment FOR learning is the heart of assessment reform in the Region and Hong Kong is no exception to this Regional trend. The Basic Competency Assessment (BCA) is designed by the Hong Kong SAR Government to be a formidable vehicle to drive the local assessment reform.

# 2. The Basic Competency Assessment in Hong Kong

## 2.1 Components of the Basic Competency Assessment

In 2000, the Education Bureau (then Education and Manpower Bureau) of the Hong Kong SAR Government accepted invitation by the Education Commission to design, develop and implement a Basic Competency Assessment (BCA) for Hong Kong. The Hong Kong Examinations and Assessment Authority (hereafter HKEAA) was subsequently commissioned by the Education Bureau in 2001 to develop and administer the BCA. There are two components of the BCA, namely the BCA Student Assessment and the Territory-wide System Assessment (TSA). The BCA Student Assessment was implemented with a prototype in 2002 and full implementation in 2005. The BCA Student Assessment is a low-stake, criterion-referenced, voluntary online testing programme provided by the Hong Kong SAR Government as an option of internal assessment by schools. The curricula covered by the BCA are the subjects of Chinese, English and Mathematics from Key Stage One to Key Stage Three. Complementary remedial packages and teacher support materials are also made available by the Education Department to help teachers in designing individualized learning for students. Services provided by the BCA Student Assessment are available to all local primary schools since 2003 and to secondary schools since 2005.

The BCA Student Assessment is designed to facilitate assessment for learning (Education Commission, 2000). As the name implies, the BCA Student Assessment covers only the basic and essential elements of the curriculum and represents only the knowledge and skills which students should master for effective progress to the next higher stage of study without meeting any serious difficulty. It does not cover the full curriculum. Further, the design of the BCA Student Assessment is to be used in combination with other internal assessment methods used by teachers rather than to replace these other methods. Teachers have full autonomy to decide as to when and how best to use the BCA Student Assessment to support learning and teaching in their classes. Records and data generated from the BCA Student Assessment are kept by the HKEAA and schools can also keep a copy. The BCA Student Assessment is used for continuous improvement in learning and teaching within schools. This is in line with the "Formative Purpose" of assessment in the Education Commission Report Number 4 (2000), which is stated as follows:

> *"An important purpose of assessment is to provide information on students' strengths and weaknesses to teachers. This enables them to plan their future teaching so that it can build upon the strengths and address the weaknesses. With this information, students can decide what to concentrate on to help them meet their future educational or employment requirements. Parents may also use the information to support and guide their children in their studies.* (Education Commission, 2000, Section 5.3.3, p. 64)".

In contrast with the formative purpose of the BCA Student Assessment, the Territory-wide System Assessment (TSA) is a secure test administered centrally by the government. The purpose of TSA is to provide data for monitoring of standards and evidence-based policymaking at the system and school management levels. All schools in Hong Kong take part in TSA at designated dates. Moreover, participation of schools is compulsory in order to fulfill the system level quality assurance purpose. Different from the online delivery mode of the BCA Student Assessment, the TSA is mainly conducted on paper-and-pen format, except for the oral assessment component for Chinese and English languages. TSA is administered at the exit level of Key Stage One (Primary 3; since 2004), Key Stage Two (Primary 6; since 2005), and Key Stage Three (Secondary 3; since 2006). For individual students, TSA is low-stake in nature as the assessment aims to generate feedback for school management regarding students' standards in Chinese Language, English Language

and Mathematics for school continuous improvement. Similar to BCA Student Assessment, TSA only targets basic competencies in the three curriculum subjects. In fact, items used in TSA each year are placed in the item bank of the BCA Student Assessment in subsequent year. Essentially, items for the two components of the BCA are "identical" other than the time and format of administration.

One main objective of TSA is to provide individual schools with assessment data of the school against the reference framework of results of all schools in the same assessment in order to facilitate schools in formulating improvement strategies for effective teaching and learning. Although the formative function of TSA has been stressed by the government in all public documents, the TSA does carry with it a parallel summative function highlighted in Education Commission Report Number 4:

> "…to provide a clear and full description of what a student has achieved at the end of a course of study or stage of education. This information is useful in seeing how individual students or schools in general are performing at different stages in the education process. A cumulative record of achievement of each student may be kept to chart his or her progress through primary and secondary school (Education Commission, 2000, Section 5.3.3, p. 64)".

## 2.2 Design of the Basic Competency Assessment Student Assessment System

The BCA Student Assessment is designed as a platform to enable students, teachers and parents to understand student' learning progress in order to provide appropriate support. The design can be represented schematically in Figure 1. The BCA Student Assessment system has two components, namely, (1) the item bank and assessment bank; and (2) the online assessment and reporting platform. Items in the item bank are designed according to basic competency framework and classified according to basic competence descriptors (e.g. the competence "Using simple strategies to extract meaning from short texts" in the knowledge dimension of reading skills). The items have been pilot tested or used in previous TSA and their difficulties have been estimated using Rasch model (Rasch, 1960/1980). Consequently, teachers can make use of information on item difficulty and targeted basic competence to assemble items and create an assessment for their students.

Once created, the assessment can be stored in the BCA Student Assessment system for any future retrieval and administration, or used as a template for designing new assessments. In addition, the system has already installed a number of pre-defined assessments for each subject which can be downloaded or administered online for their students (HKEAA, 2009).

To administer an assessment already prepared using one of the methods mentioned, the teacher needs to enroll his/her students to the BCA Student Assessment system, and specify the date and time of the system. Students can take the assessment within the specified period. After the assessment, with the exception of speaking and writing tasks, students' responses are scored automatically online and feedback is provided immediately. Individual students' raw scores and the ability levels estimated from the Rasch model (Rasch, 1960/1980) are stored in BCA Student Assessment database, although only the raw scores are reported to students. Each student can store, retrieve and manage his/her own information. In this way, each student has a clear picture of his/her own progress and areas needing further work, thus giving clear direction for maximising his/her potentials. Each teacher can store, retrieve and manage assessment data of his/her own class in the BCA database for evidence-based decision making on how to better align the teaching with the achievement level of his/her students. Likewise, school administrators can store, retrieve and manage assessment data of the whole school, thus enabling evidence-based policymaking. Users Guides for teachers and students are available on the web (http://www.hkbca.edu.hk/eng/doc/user_guide/teacher_2005_eng.pdf).

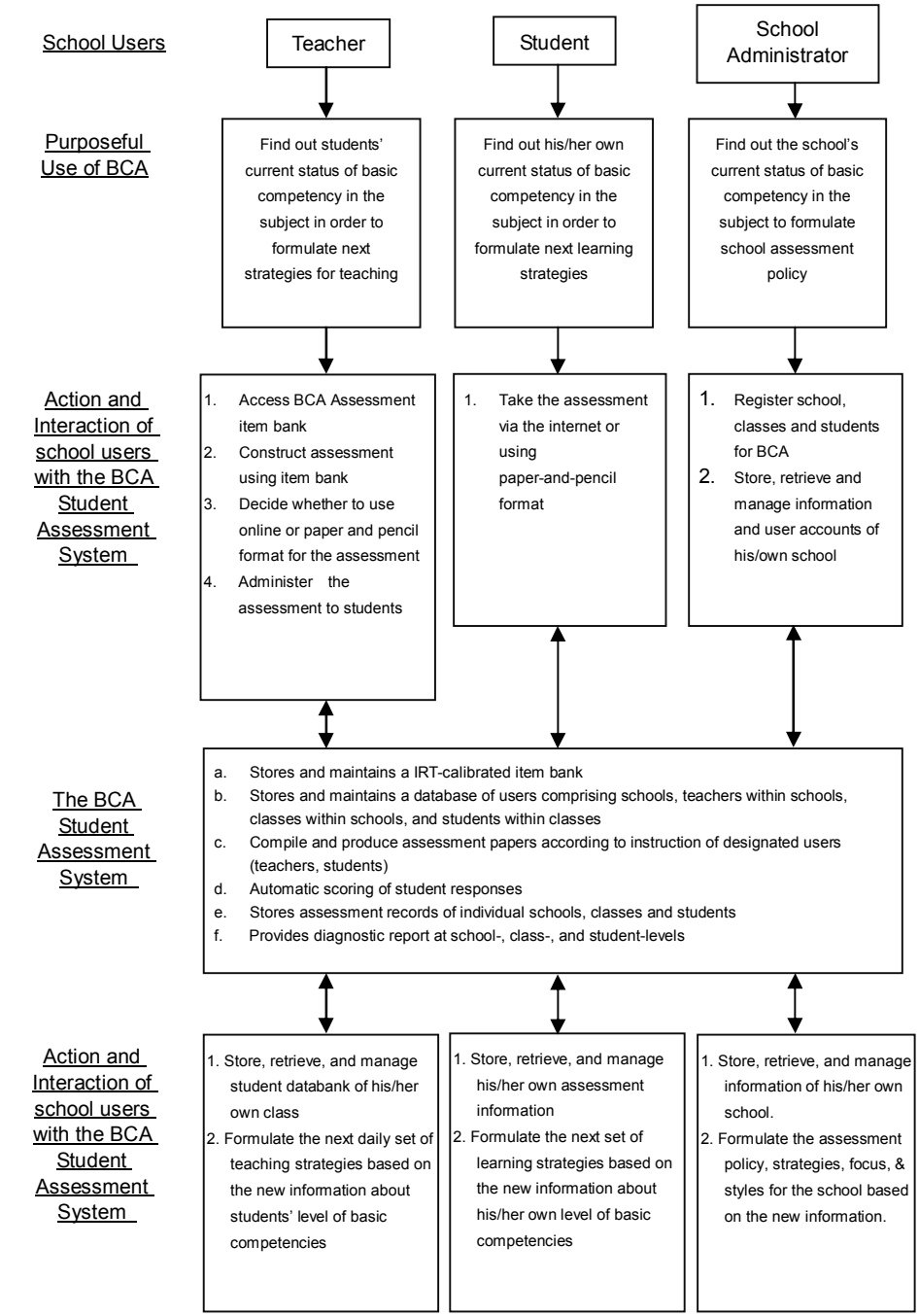
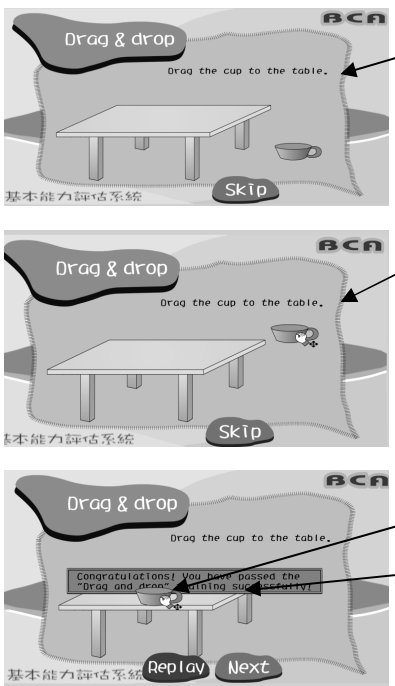



Figure 1 Design of BCA Student Assessment

## 2.3 Design of the Basic Competency Assessment Student Assessments and Items

The current BCA Student Assessment contains items covering the key learning areas of Chinese, English and Mathematics at all levels from Primary 1 to Secondary 3. The item formats take on a variety of formats and supported by multi-media technology. Response formats include choosing the answer by clicking, highlighting, dragging and dropping, dropping down a list, or selecting one or more answer boxes. A demo site is available at http://www.hkbca.edu.hk/eng/swf/demokit.htm (Figure 2) to get teachers and students familiarize with the response and item formats.




2. There is audio as well as written instruction from the computer inviting the user to "Drag the cup to the table".

1. The user
   - Place the cursor over the cup
   - hold down the right button of the mouse
   - drag the cup to the table

3. After the user has successfully dragged the cup to the table and released the mouse, there is a congratulation statement to encourage the user on learning how to use the "Drag & drop" tool.

Figure 2 Demonstration of the Drag & drop response skill for online BCA

Source: Hong Kong Examinations and Assessment Authority (n.d.a)

The BCA Student Assessment for each subject is curriculum-based. Assessment of the English Language subject focuses on the four language skills of listening,

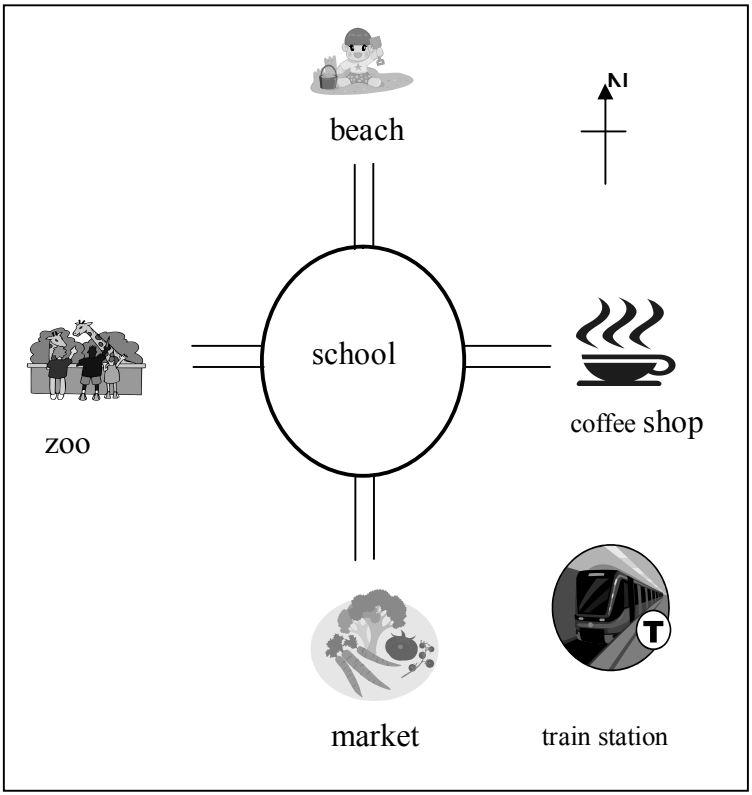
reading, speaking and writing. Listening and reading are assessed online. Each skill is further divided into basic competencies. For instance, competencies for writing in English for Primary 3 students include the ability to:

- Write and/or respond to short and simple texts with relevant information and ideas with the help of cues; and
- Write short and simple texts using a small range of vocabulary, sentence patterns and cohesive devices fairly appropriately with the help of cues despite some spelling and grammatical mistakes (HKEAA).

A number of assessment items are stored in the BCA Student Assessment databank for each of these competencies.

An item similar to one used in the BCA Student Assessment is displayed in Figure 3. The figure presents the location of a school relative to shops and community facilities. Students are asked to locate the directions from reading the figure. Three items are asked of students. The first part of the item is a straight forward one on direction, "The zoo is to the _____ of the school". The second part of the item, "Mary can go _____ to the coffee shop from the school, and then go _____ to the train station" requires competency to combine two pieces of information about the relative positions of three objects. The third part of the item, "The market is at the _____ of the beach" requires the student to read the relative direction of two objects (market and beach) by reading beyond the obstructing object (school) in between. It can be seen from this example that the BCA Student Assessment items make use of concrete situations to decipher students' level of knowledge. Items are contextualized in situations familiar to the students (e.g. school, zoo, market place, etc.). This is different from traditional examination where the items tend to be more abstract and alien from students' everyday lives.

Q8. The following are the community facilities around a school.




Please write the direction in the blank space:

(a) The zoo is to the _____ of the school.
(b) Mary can go _____ to the coffee shop from the school, and then go _____ to the train station.
(c) The market is at the _____ of the beach.

Figure 3 An item similar to one used in BCA Student Assessment

## 2.4 Reporting of the Basic Competency Assessment Student Assessment

All responses to the BCA Student Assessment, with the exception of writing and speaking items, are automatically marked by the BCA system. The students may retrieve their own report and teachers may access each student's report as well as the

report of the whole class. School administers can in addition access the report of the whole school. Presented in Figure 4 is an example of student report generated from the BCA Student Assessment. It can be seen that the student's answer is listed against the correct answer in the report. In addition, the answer is scored either right or wrong, and if the answer is wrong possible mistakes and misconceptions are listed. The Basic Competency which the item is assessing is also listed in the same report. As such, the report is information rich for the student and his/her teacher to take follow up actions in teaching and learning.

| 2009-07-29 Mathematics Assessment Construction | | | | | |
|---|---|---|---|---|---|
| Student Name | SECONDARY SCHOOL STUDENT ONE | Student Number | 10003841 | | |
| Class | 1A | Subject | Mathematics | | |
| Date | 2009-07-29 | | | | |
| Number of BCs | 4 | | | | |
| Item Number | Student's Answer | Correct Answer | Right/Wrong | Possible mistakes and misconceptions | Basic Competency |
| 1 | 76 | 75 | ☒ | • Not able to calculate the weighted mean of a set of data | • 3 |
| 2 | 14 | 15.1 | ☒ | • Not able to calculate the mean from a set of ungrouped data | • 1 |
| 3 | 7400 | 7700 | ☒ | • Not able to calculate the median from a set of ungrouped data | |
| 4 | 4 | 3.3 | ☒ | • Confused the arithmetic mean | • 2 |
| ✂ | (skipped…) | | | | |
| 14 | 54 | 73 | ☒ | • Not able to calculate the weighted mean | • 3 |
| 15 | C | C | ☑ | | • 4 |
| | Total of correct answer | | 8 | | • |

Figure 4 Example of student report from BCA Student Assessment

Source: Hong Kong Examinations and Assessment Authority (n.d.b)

The class report (Figure 5) is a summary table of the number of students in the class getting a correct answer for each item in the assessment. Each row of the table

contains the responses of one student and each column contains the responses of all students to an item. The rows and columns are sorted in descending order according to the row total and column total respectively. In this arrangement, items in the table are arranged from the easiest on the left to the hardest on the right, and students in the class are arranged from the most able on the top to the least able at the bottom. In addition, each cell in the table is colour-coded with green cell representing a correct answer and colourless cell representing a wrong answer. The arrangement of the table facilitates the teacher to identify quickly areas of strength and weaknesses for each student, and implement remedial actions where appropriate. The arrangement has satisfied the prior conditions for the student-problem chart (Harnisch, 1983; Sato, 1985; 余民寧, 1998) analysis to be undertaken. It would be beneficial for the student-problem chart to be included for use by teachers.

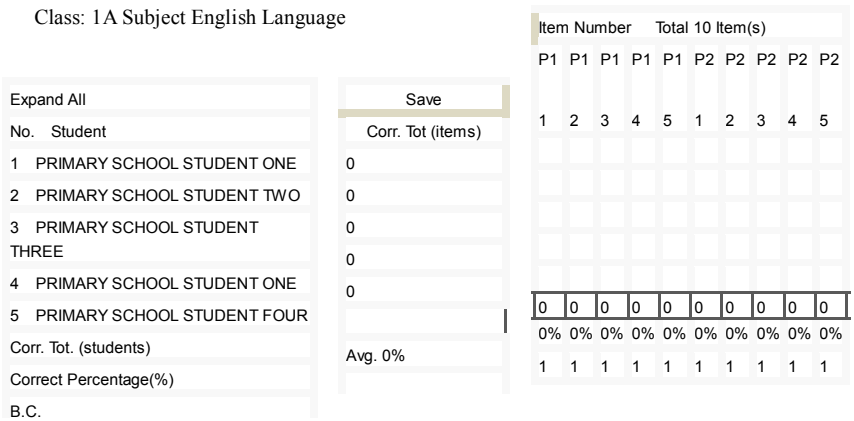



Figure 5 Example of class report from BCA Student Assessment

Source: Hong Kong Examinations and Assessment Authority (n.d.b)

The missing components are the student- and problem- curves which indicate the expected number of items correct for each student, and the expected number of students correct for each item respectively. The other missing component is the Modified Caution Index (Harnisch, 1983; Sato, 1985; 余民寧, 1998) which is a number ranging from zero to one. Modified Caution Indices of values greater than 0.5 suggest that the pattern of response may be deviated from the expected pattern

and the teacher should pay attention to the student having the discrepancies because such discrepancies indicate that the student lacks readiness, does not have examination skills, is careless, has cheated, etc. Modified Caution Index greater than 0.75 suggests that there are major problems in the pattern of responses of the student (Sato, 1985) and the teacher should be seriously concerned.

# 3. Territory-wide Systems Assessment in Hong Kong

## 3.1 Design of the Territory-wide Systems Assessment

The Territory-wide Systems Assessment (TSA) is a standard-referenced low-stake assessment in English, Chinese and Mathematics conducted at the end of Key Stages One (Primary 3), Two (Primary 6), and Three (Secondary 3). Each year, with the exception of oral assessment, all students at Primary 3, Primary 6 and Secondary 3 from all schools participate at the TSA on a date (usually in June) designated by the government. For small primary and secondary schools, 12 students are randomly selected for the oral assessment of English Language on the day of assessment and another 12 students for oral assessment of Chinese Language. Twenty-four students are selected for the oral assessment of each language for large schools. The TSA involves about 210,000 students at Primary 3, Primary 6 and Secondary 3 from about 620 primary and 450 secondary schools each year (HKEAA, 2008, p. 9).

Assessment items for each subject are developed each year by a panel of experienced subject teachers, subject experts and assessment experts from local schools, the Curriculum Development Institute of the Education Bureau, the HKEAA, and universities. Each panel drew up a test blueprint according to a set of basic competency descriptors and the Curriculum Guide of the Curriculum Development Council (HKEAA, 2005), focusing on basic competencies in the subjects required for students to progress to the next higher level without major difficulties. The TSA assessment items are developed according to the blueprint, which ensures that different contexts, text types and item types are adequately represented in the assessment. The panel reviews and endorses the items through several rounds of discussion and consideration on each item. The process of assessment development of the TSA is represented schematically in Figure 6.

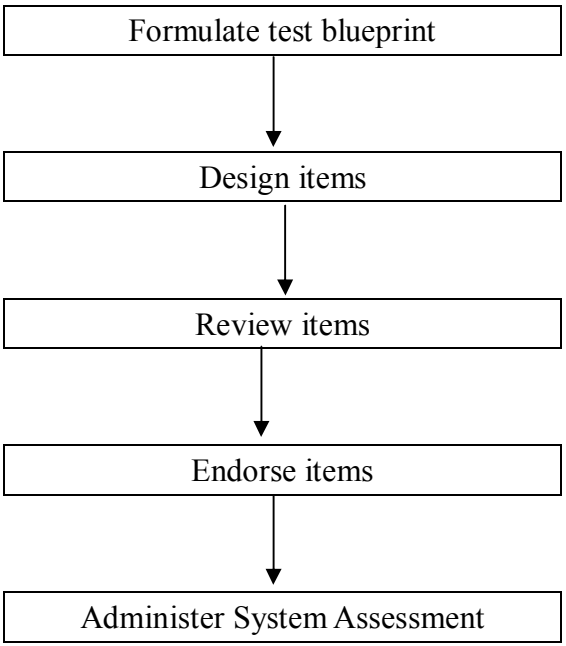



Figure 6 Process in developing the TSA

Source: Hong Kong Examinations and Assessment Authority (2008, p. 3)

The TSA is mainly conducted in the written format, but it also includes the oral assessments of the two Languages and Chinese Audio-visual assessment at the end of Key Stages One to Three. It can be seen from Table 1 that the written papers for both English and Chinese include Listening, Reading and Writing for Key Stages One, Two and Three. On the other hand, the speaking assessments change from Key Stage One to Key Stage Three to take into consideration language developments of children across these Key Stages. At Key Stage One, Storytelling and Group Interaction in Chinese are used to assess the Chinese speaking competency of children. Storytelling is also used at Key Stage Two, but at Key Stage Three, Presentation and Group Discussion are used in speaking assessment. Similar approach has been used for the assessment of English speaking, with Reading Aloud only used at Key Stages One and Two, and Group Interaction only at Key Stage Three. The method of speaking assessment is selected to align with students' language developments in order to enhance the validity and utility of the assessments. For the subject of mathematics, the competency area of Algebra is

introduced in Key Stage Two. It is combined with Number to form one competency area, while Measures, Shape and Space which are separately assessed in the earlier Key Stages are combined as one area in Key Stage Three.

Table 1 Number of items in each sub-paper (2008)

| Subject | Basic Competency Area | Number of Sub-Papers | | |
|---|---|---|---|---|
| | | Key Stage One | Key Stage Two | Key Stage Three |
| **Chinese Language** | | | | |
| Written Paper | | | | |
| | Listening | 2 | 2 | 2 |
| | Reading | 4 | 4 | 3 |
| | Writing | 2 | 4 | 3 |
| | Audio-visual | 2 | 2 | 1 |
| Speaking | | | | |
| | Storytelling | 4 | 4 | N/A |
| | Group Interaction | 4 | N/A | N/A |
| | Presentation | N/A | 4 | 4 |
| | Group Discussion | N/A | 4 | 4 |
| **English Language** | | | | |
| Written Paper | | | | |
| | Listening | 4 | 4 | 3 |
| | Reading | 4 | 4 | 3 |
| | Writing | 4 | 4 | 3 |
| Speaking | | | | |
| | Reading Aloud | 4 | 4 | N/A |
| | Personal Experience | 4 | N/A | N/A |
| | Picture Description | 4 | N/A | N/A |
| | Teacher-Student Interaction | N/A | 4 | N/A |
| | Presentation | N/A | 4 | 4 |
| | Group Interaction | N/A | N/A | 4 |
| **Mathematics** | | | | |
| Written Paper | | | | |
| | Number | 4 | 4 | N/A |
| | Measures | 4 | 4 | N/A |
| | Shape and Space | 4 | 4 | N/A |
| | Data Handling | 4 | 4 | 4 |
| | Algebra | N/A | 4 | N/A |
| | Number & Algebra | N/A | N/A | 4 |
| | Measures, Shape & Space | N/A | N/A | 4 |

Note: Areas where the Basic Competency is not assessed (not applicable) is marked with N/A.

Given that there are many basic competencies to be included in the assessment, the items are arranged in a number of assessment papers (booklets) using the method of multiple matrix sampling (Hutchison, Kendall, Bartholomew, Knott, Galbraith, & Piccoli, 2000) whereby each student attempts a portion of the items and together all the items are taken by some students across the assessment papers. The number of papers varies according to subject and Key Stage, ranging from 2 papers in Chinese Language (Audio-visual at Key Stage 3) to 4 papers in Mathematics (all Key Stages) in 2008 (HKEAA, 2008). Details are presented in Table 1.

The panel of each subject decides on the number of items, the number of assessment papers, and the duration of each paper for each Key Stage. It was decided that each student attempts one paper of each subject. Each paper ranges from 40 minutes to 140 minutes in duration (Table 2). Items developed are reviewed and endorsed by the panel to ensure quality. The final set of items is assembled into papers for each subject. Some items are used in more than one paper as linkage items for equating of test scores across papers. For instance, in 2008, 45 items have been developed to assess basic competencies in "Number" of Key Stage One students and each paper has included between 14 and 16 of the items (HKEAA, 2008). Linkages across papers are strong for each subject at each Key Stage. When individual competencies are considered, linkages are stronger in some competency areas than others. In general, reading assessments of the two languages, which have at least 10 items used in more than one paper, have stronger links across papers than speaking assessments, which do not have any common item across papers, for all Key Stages. For mathematics, assessment of number has stronger linkage than assessment of shapes. Nevertheless, information on exact number of common items across papers for each competency area is not available.

Table 2 Duration of each sub-paper (2008)

| | Duration (minutes) of Each Sub-Paper | | |
|---|---|---|---|
| Subject | Key Stage One | Key Stage Two | Key Stage Three |
| Chinese Language | 85 min | 115 min | 140 min |
| English Language | 45 min | 80 min | 90 min |
| Mathematics | 40 min | 50 min | 65 min |

An example assessment is taken from the Territory-wide System Assessment 2005 for English reading and writing at Primary 6 (Key Stage II), of which 55 minutes are allowed. The assessment is constructed around the theme of libraries. It has six Parts, each of which may be further divided into two sub sections (e.g. Part 1A, Part 1B). For Part 1, students are presented with two notices on the opening hours of the Central Library and the Mobile Library in Lam Tin District at different days of the week, then seven comprehension questions in the form of multiple choice items are asked in Part 1A of the assessment. Their answers can be found from inferences or literal meaning from the text. Some are coherence and others cohesion questions. An example item is, "The Mobile Library will NOT be open from…". Students have to choose from four options of non-opening hours. Part 1B of the assessment comprises four questions requiring short answers. For example, "Which day has the shortest opening hour for the Central Library?" The design of Part 2 is similar, although the genre of the stimulating material is that of a poem. The stimulating material of Part 3 is a riddle about nature and only short answers are required. Part 4 consists of a notice in the library about competition and students have to respond to multiple choice items with information derived from contents of the notice. Students have also to complete an envelope. Part 4B and Part 4C continues to surround the theme of the competition, with assessment tasks including completing a form to enter the competition, constructing rules for the competition. Part 5 of the assessment concerns choosing books to read from the library. Two pieces of stimulus materials about the book covers are presented and students have to answer questions derived from their understanding what has been said about the two books from the book covers. The stimulus material for Part 6 is a letter, and students have to write a letter in response to the stimulus letter.

As can be seen from the above example, the TSA items for English reading and writing at Primary 6 assess students' comprehension of text written in different genre; contents of the text are related to the daily lives of Primary 6 students in Hong Kong; the items require different skills including comprehension, construction, evaluation, analysis and critique, all embedded in the context of two libraries in Hong Kong. These features are different from traditional assessment where the items were often de-contextualised and different items in the same assessment usually covered different themes.

It is unfortunate that the reliability and validity data of the TSA is confidential information of Hong Kong, owned by the Hong Kong Education Department. Otherwise it would be appropriate to include such information here.

## 3.2 Standard Setting and Standard Maintenance of the Territory-wide Systems Assessment

The standard setting procedures make use of a methods base on a combination of the modified Angoff method (1971), the Bookmark method (Karantonis, & Sireci, 2006) and Rasch measurement (Rasch, 1960/1980). The standard setting process has involved a team of local and overseas experts in assessment and measurement consultants between 2003 and 2004 for expert support, such that the standards are defensible, fair, valid, reliable and relevant to the Hong Kong curriculum. Importantly, the process in undertaken by a panel of judges and has representations of local practitioners, assessment experts and subject experts from schools, the government as well as the tertiary sector. Each subject of English, Mathematics and Chinese has two standard setting groups, each of 24 members of whom 20 are subject teachers who have at least four years of teaching experience in the subject, have had ample experience in item setting and are familiar with the items. The other members are officers of the Curriculum Development Institute, Education Bureau, and subject panel chair of the HKEAA. The panels are chaired by academics from local universities. The subject teachers are selected from each of the three school bandings, representing high-, medium- and low-achieving schools. At the end of the standard setting process, Item Response Modelling methods are used to identify those judges whose ratings form outliers (which means extremely stringent or extremely lenient compared to the other judges in the panel), or unstable (which means sometimes very stringent and other times very lenient in the process). After moderating such outlier cases, standards are formed by combining statistically cut scores from the two independent panels (HKEAA, 2005).

Equating using the common person - common items method is used to maintain standards across years. A secure set of items called the Research Test is administered to a sample of students in year 1 (e.g. 2007) and again in the consecutive year 2 (e.g. 2008). That is, the 2007 Equating Sample takes the Research Test and TSA-2007, and the 2008 Equating Sample takes the Research Test and TSA-2008. Scores of the Equating Sample in the Research Test are then used as anchor to link the TSA assessments in Years 2007 and 2008. Using separate calibration across adjacent years (as opposed to concurrent calibration across all years), the 2008 TSA scores are equated with previous years. After equating, cut scores used in previous years are used for data in Year 2008 to calculate the percentage of students who have reached basic level of competencies.

Presented in Table 3 are summary statistics on the percentages of students achieving basic competency between 2004 and 2008 in the three subject areas for students at the exit of Key Stages One, Two and Three. A number of observations can be made. First, there is strong consistency in standards across the five years in all Key Stages as shown by the small range and standard deviation across years for each subject at each Key Stage, especially for Key Stages Two and Three for the Languages and Key Stage Two for Mathematics. Second, the percentage of students achieving basic competency decreases, albeit only slightly (by less than 10%), from end of Key Stage One to end of Key Stage Three for each of the three subjects. Third, the percentage of students achieving basic competency is most in Mathematics, followed by Chinese, and then English (Figure 7). Further, at the end of Key Stages, between 15% (Key Stage One; around 10,000 pupils) and 30% (Key Stage Three; around 20,000 adolescents) of students who have not reached basic competency and that means a large number of children each year. Unfortunately, the error estimates of these percentages of students reaching/not reaching standards are confidential information and cannot be disclosed by the Hong Kong Examinations and Assessment Authority. Consequently, no error bars cannot be inserted in the bar graph in Figure 7.

Table 3 Percentage of students achieving basic competency, 2004 - 2008

| Subject | Key Stage | Min | Max | Median | Mean | SD |
|---|---|---|---|---|---|---|
| Chinese Language | One (End of P3) | 82.70 | 85.40 | 84.90 | 84.58 | 1.08 |
| | Two (End of P6) | 75.80 | 76.70 | 76.45 | 76.35 | 0.39 |
| | Three (End of S3) | 75.60 | 76.50 | 76.20 | 76.10 | 0.46 |
| English Language | One (End of P3) | 75.90 | 79.50 | 79.30 | 78.58 | 1.52 |
| | Two (End of P6) | 70.50 | 71.50 | 71.30 | 71.15 | 0.44 |
| | Three (End of S3) | 68.60 | 69.20 | 68.90 | 68.90 | 0.30 |
| Math | One (End of P3) | 84.90 | 86.90 | 86.90 | 86.48 | 0.88 |
| | Two (End of P6) | 83.00 | 84.10 | 83.80 | 83.68 | 0.47 |
| | Three (End of S3) | 78.40 | 79.90 | 79.80 | 79.37 | 0.84 |

Notes:

1.Chinese and English Language results include TSA results in Listening, Reading and Writing;

2.Chinese Audio-visual component included in the calculation of the cut score at the S.3 level in 2007 and 2008

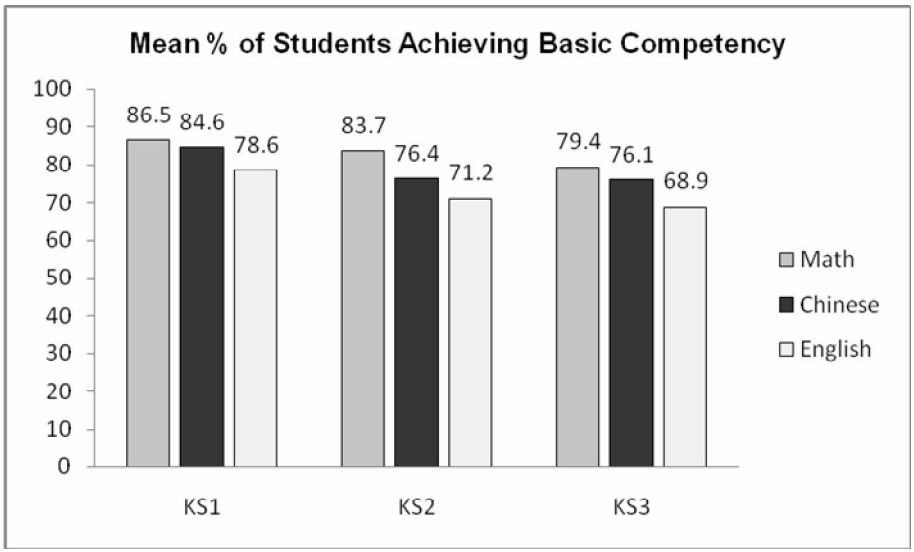



Figure 7 Mean percentage of students achieving basic competency between 2004 and 2008

## 3.3. Marking of Territory-wide Systems Assessment

Objective items are automatically scored by the computer. Experienced language teachers are invited to be Oral Examiners after training by the HKEAA. The most innovative about TSA operation however concerns Onscreen Marking, introduced to the HKEAA in 2007 and adopted in 2008 for the marking of TSA written assessments. Onscreen Marking is a fast and flexible marking and script management process that minimizes the chance of human error. It enables effective monitoring of examiner reliability, allows problems to be identified and interventions to be implemented at an early stage, and provides instant and detailed feedback to marking-centre supervisors about examiner and student performance. Importantly, on-screen marking has eliminated the problem of loss of examination scripts, a problem that has plagued Hong Kong high-stake examination for years when in the olden days examiners were allowed to take scripts home for marking. Onscreen Marking has won Hong Kong Examination and Assessment Authority several prestigious awards including (1)MIS 2008 IT Excellence Awards 2008 - Best Change Management (Government); (2)HKICT Awards 2007 – Best Business (Application) Certificate of Merit organized by the Hong Kong Computer Society; and (3)Government Technology Awards 07 - Runner up.

## 3.4. Reports provided by the Territory-wide Systems Assessment

Reports of the TSA are at school-level only. Results at individual student level are not reported. Each school is provided with two TSA reports on the school's own results each year, namely the confidential school report, and item analysis reports. A reference framework of individual school results mapped against the territory-wide basic competency standards in Chinese Language, English Language and Mathematics at the end of the three Key Stages is provided to facilitate interpretation. In addition, three TSA 'Supplementary Reports' are given to schools. The first two of these reports exclude the performance of students with mild intellectual disabilities and students with different/special learning needs (HKEAA, 2008, p. 16). The third 'Supplementary Report' provides longitudinal results tracking growth of students in the same school from Primary 3 (in 2005) to Primary 6 (in 2008)(ibid, p. 16). Schools have to observe confidentiality of school results on the TSA. They are not allowed to make public these results. Reports of the performance of all schools in Hong Kong are available to the general public and the government for policy review and formulation.

## 3.5 Public Acceptance of TSA

The Hong Kong SAR government has invested huge resources in introducing the Basic Competency Assessment. In order to evaluate the acceptance of TSA by local practitioners, the HKEAA has undertaken a survey in 2008 on the perceived usefulness of the TSA reports by schools and teachers, and how the reports are being used. The survey was sent to 1081 (624 primary and 457 secondary) schools, with a return rate of 67.3%, which meant 727 schools (432 primary and 293 secondary). The study found that schools were satisfied with the TSA report. Over 90% of primary and secondary respondents considered 'Subject Results and Student Performances' in the TSA report to be the most valuable. An overwhelming majority of teachers (97% primary and 94% secondary) indicated that they changed their teaching strategies in response to the TSA data. Further, about 16% respondents suggested HKEAA to provide more information about student reports, advice on teaching and learning strategies, and comparison with other like schools in the district region or banding. These results suggest that the TSA has had significant positive impact on the teaching and learning of Hong Kong schools.

# 4. Conclusion

This article presents an overview of the assessment of basic competence in Hong Kong. The BCA has taken a two-prong approach in reforming Hong Kong assessment. First, the BCA Student Assessment has provided Hong Kong schools a strong platform of assessment for learning in the daily praxis of Hong Kong teachers. Second, the annual Territory-wide System Assessment is designed to provide schools and the government system level data on the performance of students at exit of Key Stages One to Three. Since its implementation, the BCA Student Assessment has received warm support from schools. Then the survey conducted by HKEAA showed that schools and teachers found reports generated from the TSA helpful to their teaching and learning. It seems that the BCA has gained leverage against the firmly embedded examination-oriented culture (Pong & Chow, 2002).

The BCA is a timely initiative within the international context of assessment for learning movement as well as within the local context of structural reform in the education system. Hong Kong's education system has moved from seven years of secondary education to six years of secondary education starting from the 2009 school year. Associated with this structural change are two major changes in its public examination systems. First, the current Hong Kong Certificate of Education Examination, taken by students at the end of Secondary 5, and Hong Kong Advanced-level Examination, taken by students at the exit of Secondary 7, are to be replaced by one examination, namely the Hong Kong Diploma of School Education. Second, the norm referenced public examinations are to be replaced by a standards-based reporting system. The standards-reference reporting system of the TSA report is a good precursor to these changes in the public examination system and research by the HKEAA suggests that the initiatives are welcome by Hong Kong school sector.

The Territory-wide Systems Assessment (TSA) in Hong Kong, successfully developed and implemented by the HKEAA is an initiative similar to national testing elsewhere and is an important milestone in the assessment reform of Hong Kong. The TSA not only provides policymakers with valid and reliable data on system-wide level of competence in Chinese, English and Mathematics of students at Primary 3, Primary 6, and Secondary 3 levels, it also signifies a giant step of reforming Hong Kong education from the "assessment for learning" tradition to an era where "assessment for learning" is reality in practice. At the time of writing, the HKEAA is pilot testing its Computer Scoring and Delivery System for the oral

assessment of the TSA, with an aim to launch it in 2010-2011. Before that, the HKEAA is hoping to launch its prototype Computerised Adaptive Testing for Mathematics in 2009. Further, the HKEAA will have longitudinal data at Secondary 3 in 2011 of the first cohort who joined TSA at Primary 3 in 2004. All of these are exciting developments for Hong Kong education.

## References

American Psychological Association Board of Educational Affairs. (1997, November). *Learner-centered psychological principles: A framework for school reform and redesign*. Washington, DC: American Psychological Association.

Angoff, W. H. (1971). Scales norms and equivalent score. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed.) (pp. 508-600). Washington, DC: American Council on Education

Biggs, J. (1996). The assessment scene in Hong Kong. In J. Biggs (Ed.). *Testing: To educate or to select* (pp. 3-12). Hong Kong: Hong Kong Educational Publishing Co.

Black, P., & Wiliam, D. (1998). *Inside the black box*. London: King's College.

Cheng, Y. C., & Townsend, T. (2000). Educational change and development in the Asia-Pacific region: Trends and issues. In T. Townsend & Y. C., Cheng (Eds.), *Educational change and development in the Asia-Pacific region: Challenges for the future* (pp. 317-344). The Netherlands: Swets and Zeitlinger.

Delors, J. (1998). *Education for the twenty-first century: Issues and prospects: Contributions to the work of the International Commission on Education for the Twenty-first Century, chaired by Jacques Delors*. Paris: Unesco.

Education Commission. (1990). *Education Commission Report Number 4: The curriculum and behavioural problems in schools.* . Hong Kong: Printing Department.

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: Printing Department.

Harnisch, D. L. (1983). Item response patterns: Applications for educational practice. *Journal of Educational Measurement, 20*(2), 191-206.

Hong Kong Examinations and Assessment Authority. (2005). *Territory-wide System Assessment 2005 Report on the Basic Competencies of Students in Chinese Language, English Language and Mathematics Key Stages 1 – 3*. Hong Kong: Hong Kong Examinations and Assessment Authority.

Hong Kong Examinations and Assessment Authority. (2005). *Territory-wide System Assessment 2005 Report on the Basic Competencies of Students in Chinese Language, English Language and Mathematics Key Stages 1 – 3*. Hong Kong: Hong Kong Examinations and Assessment Authority.

Hong Kong Examinations and Assessment Authority. (2008). *Territory-wide System Assessment 2008 Report on the Basic Competencies of Students in Chinese Language, English Language and Mathematics Key Stages 1 – 3*. Hong Kong: Hong Kong Examinations and Assessment Authority.

Hong Kong Examinations and Assessment Authority. (2008). *Territory-wide System Assessment 2008 Report on the Basic Competencies of Students in Chinese Language, English Language and Mathematics Key Stages 1 – 3*. Hong Kong: Hong Kong Examinations and Assessment Authority.

Hong Kong Examinations and Assessment Authority. (2009). *Basic Competency Assessment: Student Assessment*. Retrieved 16 August, 2009, from http://www.hkbca.edu.hk/index_eng.htm

Hong Kong Examinations and Assessment Authority. (n.d.a). *BCA Demokit*. Retrieved 11 November, 2009, from http://www.bca.hkeaa.edu.hk/web/SA/demokit/demokit.htm.

Hong Kong Examinations and Assessment Authority. (n.d.b). *Reporting system*. Retrieved 11 November, 2009, from http://www.bca.hkeaa.edu.hk/web/SA/en/ReportingSystem.html

Hutchison, D., Kendall, L., Bartholomew, D., Knott, M., Galbraith, J., & Piccoli, M. (2000). Reliability of assessment of reading ability in three Countries. *Quality and Quantity, 34*(4), 353-366.

Karantonis, A., & Sireci, S. G. (2006). The bookmark standard-setting method: A literature review. *Educational Measurement, Issues and Practice, 25*(1); 4-13.

Mok, M. M. C., Gurr, D., Izawa, E., Knipprath, H., Lee, I. H., Mel, M. A., Palmer, T. Shan, W. J., & Zhang, Y. (2003). Quality assurance and school monitoring. In J. P. Keeves & R. Watanabe (Chief Eds.), *International handbook of educational research in the Asia-Pacific region* (pp. 945-958). *Kluwer International Handbooks of Education: Volume 11*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Morris, P. (1985). Teachers' perceptions of the barriers to the implementation of a pedagogic innovation: A South East Asian case study. *International Review of Education, 31*, 3-18.

Pong, W. Y., & Chow, J. C. S. (2002). On the pedagogy of examinations in Hong Kong. *Teaching and Teacher Education, 18*(2), 139-149.

Rasch, G. (1960/1980). *Probabilistic models for some intelligence and attainment tests* (expanded ed.). Chicago: The University of Chicago Press.

Rowe, K. (2005). *Evidence for the kinds of feedback data that support both student and teacher learning*. Paper presented at the 2005 Using Data to Support Learning Conference, ACT, Australia. Paper retrieved 12 August, 2009, from ACEReSearch: http://research.acer.edu.au/research conference 2005/6

Sato, T. (1985). *Introduction to student-problem curve theory analysis and evaluation*. Tokyo: Meiji Tosho.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics, 118*(4), 1279-1333.

Yeung, S. P.-M. (1996). Teaching environmental issues in school geography: The Hong Kong experience. *International Research in Geographical and Environmental Education, 5*(2), 117-130.

余民寧（1998）。S-P 表分析在教育評量上的應用。載於高雄市政府公務人力資源發展中心編，**多元教學評量**（頁 29-43）。高雄市：高雄市政府公務人力資源發展中心。

# Basic Competency Assessment in Hong Kong

**Magdalena Mo Ching Mok**

This article presents an overview of assessment reform in terms of the Basic Competency Assessment (BCA) in Hong Kong. The BCA has two components, namely, the BCA Student Assessment and the Territory-wide System Assessment (TSA). Items used in the TSA are put back into the BCA Student Assessment item bank the following year, so essentially the two components have identical item type and level. The former is designed to provide Hong Kong schools a strong platform of assessment for learning and is to be used by teachers for their daily teaching and the later is to give schools and the government system level data on student performance at Key Stages One to Three. The BCA has received very favourable support from schools and teachers. The development of the BCA is grounded on recent research findings that the feedback is one of the most important factor contributing to learning. The BCA system aims to provide quality feedback in order to help teachers to better align teaching with students' levels and to support schools and the government in formulating evidence based policymaking. This paper gives the theoretical underpinning of the BCA, a description of its components, designs, and reporting features.

Keywords: basic competency assessment, feedback, assessment for learning, assessment reform, Hong Kong

Magdalena Mo Ching Mok, Chair Professor, Hong Kong Institute of Education

主題文章