Views From the Chalkface: English Language School-Based Assessment in Hong Kong

Chris Davison
The University of Hong Kong

The Hong Kong Examinations and Assessment Authority (HKEAA) has recently moved from norm-referenced to standards-referenced assessment, including the incorporation of a substantial school-based summative oral assessment component into the compulsory English language subject in the Hong Kong Certificate of Education Examination (HKCEE). Starting in Form 4, teachers now assess their own students’ oral English language competencies through a range of classroom-embedded activities over 2 years (SBA Consultancy Team, 2005). This high-profile assessment initiative marks a significant shift in policy as well as in practice for the HKEAA. Although school-based assessment (SBA) is in line with the Education and Manpower Bureau’s general move to align assessment with curriculum reforms, in the early stage of implementation the reforms raised a number of concerns in the wider school community, including sociocultural, technical, and practical concerns. This article first describes the specific content and structure of the HKCEE English Language SBA component. It then reports on the result of the initial analysis of teachers’ and students’ responses to the initiative in the first stage of its implementation, including the perceived benefits for learning and teaching. The article concludes with a brief overview of how this initial analysis led to the development of a number of subsequent research studies aimed at monitoring and developing teacher knowledge and skills and evaluating more systematically the impact of the reform on teachers, students, and schools in Hong Kong.

The Hong Kong Examinations and Assessment Authority (HKEAA) has recently moved from norm-referenced to standards-referenced assessment, including the
incorporation of a substantial school-based summative oral assessment component into the compulsory English language subject in the Hong Kong Certificate of Education Examination (HKCEE), a high-stakes examination for all Form 4–5 (F4–5) students.\(^1\)

In school-based assessment (SBA), assessment for both formative and summative purposes is integrated into the teaching and learning process, with teachers involved at all stages of the assessment cycle, from planning the assessment programme, to identifying and/or developing appropriate assessment tasks right through to making the final judgments (see SBA Consultancy Team, 2005, for a detailed description of the activities). As assessments are conducted by the students’ own teacher in their own classroom, students are meant to play an active role in the assessment process, particularly through the use of self- and/or peer assessment used in conjunction with formative teacher feedback.

This high-profile assessment initiative, led by a team of researchers at the Faculty of Education, University of Hong Kong, in partnership with the HKEAA, marks a significant shift in policy as well as in practice for the HKEAA.\(^2\) The initiative aims to align assessment more closely with the current English language teaching syllabus (Curriculum Development Council, 1999) as well as the new outcomes-based Senior Secondary curriculum, to assess learners’ achievement in areas that cannot be easily assessed by public examinations and at the same time enhance student self-evaluation and lifelong learning. Although this is in line with the Education and Manpower Bureau’s general move to align assessment with curriculum reform (Curriculum Development Institute, 2002), in the initial process of implementing the SBA initiative a number of challenges arose. This article focuses on these challenges through an analysis of the perceptions of a range of F4 teachers and students involved in the initial introduction of the reform.\(^3\)

Studies of the impact of earlier changes in the Hong Kong external examination system in English language (e.g., Andrews, 1994; Andrews, Fullilove, & Wong, 2002; Cheng, 1998, 2005) found that changes to summative assessment did not automatically lead to improvement in learning, as the teacher and school mediated the nature of the change. Studies of the implementation of the

---

\(^1\)For a full description of the 2007 HKCEE English Language syllabus, including the external examinations, see http://web.hku.hk/~sbapro/doc/Annex%206%20Revised%202007%20HKCEE%20Eng%20Lang%20Syll.pdf

\(^2\)The research team was led by the author, Dr. Chris Davison, and Professor Liz Hamp-Lyons, Faculty of Education, University of Hong Kong.

\(^3\)Subsequent studies are supporting and monitoring the implementation of SBA over its first 2 years, as well as researching the impact of the reform on a number of schools, see <http://web.hku.hk/~sbapro/projects.html> for a full list of these studies.
Target-Oriented Curriculum in Hong Kong primary schools (e.g., Adamson & Davison, 2003; Carless, 2004; Cheung & Ng, 2000) also found assessment innovation to be severely constrained by traditional school culture and by teacher, parent, and student expectations. Studies of SBA in other subject areas in Hong Kong, such as the Teacher Assessment Scheme (Yung, 2001), also suggest that there may be wide variation in teachers’ interpretations of student performance and of their role in the assessment process.

Although SBA as an integral component of the formal senior secondary examination system is established practice in a number of educational systems internationally, including Australia, New Zealand, and the United Kingdom (Black, 2001; Black, Harrison, Lee, Marshall & Wiliam, 2003; Black & Wiliam, 2001; Sadler, 1989; Wiliam, 2001), as well as in some developing countries (Chisholm et al., 2000; Pryor & Akwesi, 1998; Pryor & Lubisi, 2002), there has been little specific research into the large-scale use of SBA in English as a second or additional language. In Asia there are embryonic attempts to develop SBA in Singapore and Malaysia as a complement to external examinations at the senior secondary level but virtually no research into the issue. In Australia, several studies of the use of large-scale criterion-referenced English as a Second Language assessment frameworks in schools (Breen et al., 1997; Davison & Williams, 2002) have revealed a great diversity in teachers’ approaches to assessment, influenced by the teachers’ prior experiences and professional development, by the assessment frameworks and scales they used, and by the reporting requirements placed on them by schools and systems. Concerns have been raised about mechanistic criterion-based approaches to SBA, which are often implemented in such a way that they undermine rather than support teachers’ classroom-embedded assessment processes (Arkoudis & O’Loughlin, 2004; Black & Wiliam, 1998; Davison, 2004; Leung, 2004b).

Research into SBA internationally is further complicated by the considerable uncertainty and disagreement around the concept and by its intrinsically teacher-mediated, coconstructed, and context-dependent nature (Black & Wiliam, 1998; Brookhart, 2003; McMillan, 2003; McNamara, 1997; Stiggins, 2001). Traditional conceptions of validity and reliability associated with the still-dominant psychometric tradition of testing are themselves a potential threat to the development of the necessarily highly contextualized and dialogic practices of SBA (Hamp-Lyons, 2006; Rea-Dickins, 2006). In a traditional exam-dominated culture, formative and summative assessment are seen as distinctly different in both form and function, and teacher and assessor roles are clearly demarcated, but in the new SBA component of the HKCEE English Language, summative assessments of the students’ speaking skills are meant to be used formatively to give constructive student feedback and improve learning. Hence, the implementation of the HKEAA English SBA initiative has both theoretical
importance and significant practical implications at the local and international level.

In this article I first describe the specific content and structure of the HKCEE English Language SBA component, then report on the result of the initial analysis of teachers’ and students’ responses to the initiative in the first stage of its implementation, including the perceived benefits for learning and teaching. The article concludes with a brief overview of how this initial analysis led to the development of a number of subsequent research studies that aim to monitor and develop teacher assessment knowledge and skills and more systematically evaluate the impact of the reform on teachers, students, and schools in Hong Kong.

THE HKCEE ENGLISH LANGUAGE SBA COMPONENT: ITS CONTENT, STRUCTURE, AND PROCESSES

The SBA component, worth 15% of the total HKCEE English mark, involves the assessment of English oral language skills based on topics and texts drawn from a programme of independent extensive reading/viewing (“texts” encompass print, video/film, fiction, and nonfiction material). At the time the SBA was introduced, students were required to choose at least four texts to read or view over the course of 2 years; keep brief notes in a logbook; and undertake a number of activities in and out of class to develop their independent reading, speaking, and thinking skills. For assessment it was suggested they participate in several interactions with classmates on a particular aspect of the text they have read/viewed, leading up to making a more formal group interaction or an individual presentation on a specific text and responding to questions from their audience. The assessment format and requirements as originally specified in the introduction to the SBA in September 2005 are summarized in Table 1.4

In terms of assessment, an important distinction is made between the two kinds of oral activities—presentation and interaction—which are characterized by distinctly different organisational and communicative strategies. An individual presentation may be quite informal, depending on task and audience, but requires comparatively long turns, hence a more explicit structure and an ability to hold the attention of the audience. In contrast, an interaction, an exchange of

---

4These initial requirements were modified slightly as a result of teacher concerns over workload, systematically documented by the SBA developers. Adjustments included reducing the number of texts from four to three and reducing the number of tasks, but the assessment focus, procedures, and criteria remain unchanged. In this article I present the initial requirements, as that was the document to which the teachers and students discussed in this study were responding (see http://web.hku.hk/~sbapro/doc/LET–2007%20CE%20SBA.pdf for the modifications that have been made since September 2005).
short turns between two or more speakers, requires less explicit structuring but more attention to turn-taking skills and planning how to initiate, maintain, and control the interaction through suggestions, questions, and expansion of ideas. Both activities, or text-types, also require the students to speak intelligibly with suitable intonation, volume, and stress, using pauses and body language such as eye contact appropriately and effectively, and to draw on a range of varied vocabulary and language patterns.

A variety of assessment tasks can be used to elicit the required kinds of oral language from students, including teacher-made tasks adapted from one of the exemplars collected from F4 and F5 teachers as part of the trial of the assessment initiative (see Appendix A for an example of one of these assessment tasks). Assessment tasks can vary in length and complexity according to a number of factors, including the communicative function, the number of people involved, the position and status of the people interacting, and the nature of the response required. This diversity of assessment tasks aims to ensure schools can provide students with appropriate, multiple, and varied opportunities to demonstrate their oral language abilities individually tailored to students’ language level and interests. For instance, in an individual presentation, the more orally proficient

---

**TABLE 1**
Initial Hong Kong Certificate of Education Examination English Language School-Based Assessment Requirements

<table>
<thead>
<tr>
<th>Requirements</th>
<th>F4</th>
<th>F5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. and type of texts to be read</td>
<td>Minimum of <strong>two</strong> texts, from two categories</td>
<td>Remaining <strong>two</strong> texts, remaining categories</td>
<td><strong>Four</strong> texts, one from each category (print fiction, print nonfiction, nonprint fiction, nonprint nonfiction)</td>
</tr>
<tr>
<td>No. and timing of assessment tasks to be undertaken</td>
<td>Minimum of <strong>two</strong> interactive tasks to be undertaken anytime during F4, must be on different texts</td>
<td>Minimum of <strong>one</strong> interactive task, <strong>one</strong> individual presentation to be undertaken anytime during F5, must be on different texts</td>
<td><strong>Four</strong> tasks on four texts, one from each category</td>
</tr>
<tr>
<td>No., %, and timing of marks to be reported</td>
<td>One mark, best mark out of the two tasks, 5% of total English mark, reported at end of F4</td>
<td>Best mark for the interaction and best mark for the presentation, 10% of total English mark, reported at end of F5</td>
<td>15% of total English mark</td>
</tr>
</tbody>
</table>

*Note. Source: Hong Kong Certificate of Education Examination (2005, p. 5).*
students can be challenged by being asked to persuade the whole class to read a particular book, whereas the less orally proficient students can be asked to describe the physical appearance of a particular character to a friend. In terms of group interaction, where each student has read different texts, the more orally proficient students can be challenged by being grouped into four and being asked to agree on which book should be set as a class reader, and the less orally proficient students can be placed in pairs and asked to find the three most important differences between their texts. Students in the same school, even the same class, may do different tasks or view different texts, so long as they all have the opportunity to produce the required type of oral language.

To ensure that the oral language produced is the student’s “best” own work, and not the result of memorisation without understanding, there are several mandatory assessment conditions (SBA Consultancy Team, 2005, pp. 7–8). First, students must be assessed by their usual English teacher, in the presence of one or more classmate(s). Second, students must be familiar with the type of task used for assessment and given sufficient opportunity to produce enough oral language to be confidently assessed. To facilitate this process, teachers are allowed to ask the students questions as appropriate to prompt or extend the range of oral language produced and/or to verify the students’ understanding of what they are saying. Third, students are not permitted to refer to extended notes nor take any notes during the assessment activity.

Students are assessed according to a set of assessment criteria, consisting of a set of descriptors at each of six levels across four domains (see Appendix B, Assessment criteria), which were developed and trialled by teachers and students from a wide range of Hong Kong schools. The domains are briefly described next.

**Domain 1: Pronunciation and Delivery**

Pronunciation comprises phonology and intonation. Phonology includes the articulation of individual sounds and sound clusters, whereas intonation refers to the flow of words with appropriate stress and rise/fall across the sentence(s). Delivery is made up of two important subaspects: voice projection and fluency. Fluency refers to the naturalness and the intelligibility of a person’s speech.

**Domain 2: Communication Strategies**

Communicative strategies involve body language, timing, and asking and answering appropriate kinds of questions. Body language includes gaze, facial expressions, head movement, and body direction—the more students rely on notes or memorized material, the weaker their body language is likely to be. Timing is important; if student takes too long for an individual presentation the audience may get bored; if the student is too brief, she or he will not be able to give enough ideas or support.
Domain 3: Vocabulary and Language Patterns

The vocabulary and language patterns domain consists of three important areas: vocabulary and language patterns (including the quantity, range, accuracy, and appropriacy), and self-correction/reformulation.

Domain 4: Ideas and Organisation

The ideas and organisation domain consists of the expression of information and ideas, the elaboration of appropriate aspects of the topic, organisation, and questioning and responding to questions. Organisation works differently in individual presentations and in group interactions. In a group interaction students share the responsibility for providing enough ideas and information to carry the dialogue forward. They need to stay focused on the topic and say something at the right time to move the conversation forward by elaborating on a point another group member has made or by bringing up a new but relevant point. This kind of organizing is much harder to do in spoken than in written language, so in F4 and F5 group interactions it is not emphasised very much. However, in an individual presentation the speaker has sole responsibility for planning what she or he will say and how, and each student is expected to have thought how to organise what he or she will say.

Within each domain each feature needs to be weighed against the others holistically to reach an overall judgment. In the same way, the levels are conceptualized not as discrete entities but rather as a continuum of development, thus it is possible to talk of a “strong 5” or a “weak 3.” An assessment record (see Appendix C) is used to provide a record of the key features of the assessment activity and help standardize the assessment process. In addition, teachers are encouraged to video- or audio-record a range of student assessments to assist with standardization and feedback, involving the students as much as possible (e.g., asking students to collect a portfolio of their oral language assessments, both formative and summative, using an MP3 player or by video-recording each other). During the class assessments, which might span a number of weeks, individual teachers at the same level (i.e., F4 or F5) are encouraged to meet informally to compare their assessments and make adjustments to their own scores as necessary. Such informal interactions give teachers the opportunity to share opinions on how to score performances and how to interpret the assessment criteria.

Near the end of the school year, there is a formal meeting of all the English teachers at each level, chaired by the SBA Coordinator in each school, to review performance samples and standardise scores. Such meetings are critical for developing agreement about what a standard means (i.e., validity, consistency in and between teacher-assessors; reliability, public accountability, and professional
collaboration/support). The adjusted marks for each student are then listed on a class record. At the end of each year there is a district-level meeting for professional sharing and further standardisation. Each SBA Coordinator is encouraged to take a range of typical and atypical individual assessment records (and the video- or audio-recordings) and the class records for sharing. Once any necessary changes are made, the performance samples are archived and the scores are submitted to the HKEAA for review. Video and audio records can be compiled on a CD-ROM for storage. Maintaining notes of all standardisation meetings and any follow up action is also encouraged so schools can show parents and the public that it has applied the SBA consistently and fairly. The HKEAA then undertakes a process of statistical moderation\(^5\) to ensure the comparability of scores across the whole Hong Kong school system.

THE RESEARCH STUDY: INITIAL PERCEPTIONS OF ASSESSMENT REFORM

Brindley (1998), in a wide-ranging study of the issues arising in the implementation of outcomes-based assessment and reporting in language learning programmes in the 1990s, identified three common types of issues and problems: what he called political issues, to do with the purposes and intended use of the assessment; technical issues, primarily to do with validity and reliability; and practical issues, to do with the means by which the assessment was put into practice. As SBA is still a very new concept for Hong Kong schools, many concerns and issues were systematically gathered from key stakeholder groups, including teachers and students, during the process of the development and initial implementation of the 2007 HKCE English Language SBA component in 2005.

There were two main stages to the collection of teacher and students’ responses to and perceptions of the SBA initiative. In the first stage in January to June 2005, prior to the actual implementation of SBA, while the assessment activities, procedures, and criteria were still being developed and trialled, data were gathered from the 66 teachers and 513 students in the 21 schools involved via questionnaires, individual and focus interviews, classroom observation, and

\(^5\)The use of statistical moderation for the SBA is controversial but seen as essential to maintain community confidence in the HKCEE. Each school’s SBA results are compared with the schoolwide result of the external oral paper, and the scores of each school (but not the scores of individual students) are adjusted if there is a marked discrepancy (see [http://www.hkeaa.edu.hk/doc/tas_ftp_doc/CE-Eng-StatModerate0610.pdf](http://www.hkeaa.edu.hk/doc/tas_ftp_doc/CE-Eng-StatModerate0610.pdf) for a more detailed description of the moderation process).
stimulated recall. Data included information about teacher and student background, existing assessment practices, perceptions and beliefs about teaching and learning, and attitudes toward the assessment reform. In a follow-up questionnaire in November 2005, 3 months after the formal introduction of SBA in schools, responses to the initiative were systematically collected from more than 173 secondary schools in Hong Kong, including both English and Chinese-medium schools with different student populations, banding levels and geographic locations ($N = 500$, response rate = 34.60%). Qualitative data from both rounds of data collection were coded and analyzed using NVivo (QRS, 2002) and the key themes and patterns identified. Triangulation, peer debriefing, and member checking were then used to test the robustness of the categories.

The remainder of this article deals with the key issues and concerns highlighted as a result of this study (the quantitative data are reported elsewhere; see Davison & Hamp-Lyons, in press). Adapting Brindley’s (1998) taxonomy of factors, the issues and concerns were broadly classified into three types: sociocultural, technical, and practical. I briefly describe these in turn.

Sociocultural Issues

Brindley (1998) observed that

if the theoretical underpinnings of the (assessment) statements or the testing formats used are seen to be at variance with the strongly held views of powerful interest groups representing particular theoretical or pedagogical orientations, then their validity may well be publicly challenged, thus greatly reducing the likelihood of their adoption by practitioners. (p. 62)

Until very recently, assessment practices in Hong Kong were driven by the need to provide data to select students for higher education or employment (Biggs, 1995), hence external examination results have traditionally been the dominant way students and schools (and teachers) have been evaluated and held accountable, although this may gradually change with the recent introduction of value-added indices\(^6\) to compare schools more equitably. As in many

---

\(^6\)The value-added approach involves comparing the performance observed with performance expected of students in the school, using a regression module that examines the performance of all schools at two time points, for example, at entry to using the F1 attainment scores and at the end of HKCE using the HKCE grades. A school is classified as “improving” (i.e., it has a positive value-added score) if the observed English performance of students exceeds the expected English performance by a sufficiently large amount. Similarly, a school is classified as “declining” (i.e., it has a negative value-added score) if its students fail to achieve their expected English performance by a sufficiently large amount.
other countries in the region (e.g., Cheah, 1998), the traditional role of assessment in the classroom (as opposed to *classroom-based* assessment) has been exam preparation.

However, the official adoption of the UK Assessment Reform Group’s (1999) distinction between assessment *for* learning, and assessment *of* learning by the Education and Manpower Bureau has stimulated the beginnings of a major paradigm shift from a culture of testing to a learning and assessment culture (see Hamp-Lyons, 1999, 2006, for a more detailed discussion of this shift in the Hong Kong context). This shift, in many ways in the opposite direction to most English-speaking countries, is being accelerated by the introduction of SBA. However, extensive experience with earlier educational reforms in Hong Kong has shown that assessment theories from the international research literature cannot be incorporated into public policy without resolving the fundamental opposition, even competition, with local “cultures” and institutional discourses (Adamson & Davison, 2003; Carless, 1999). In Hong Kong cultural assumptions about assessment are deeply entrenched in the wider community, going well beyond Brindley’s (1998) “political” issues to embrace very different preconceptions about teaching, learning, and the purpose of education. This was strongly exemplified in this study by a widespread concern among teachers and students (and their parents) about the purposes of assessment, encapsulated by ongoing and public debate over the fairness of SBA. Fairness is fundamentally a sociocultural, rather than a technical, issue, “a justice that goes beyond acting in agreed upon ways and seeks to look at justice of the arrangements leading up to and resulting from those actions” (Stobart, 2005, p. 1). In a highly competitive examination-driven school system such as Hong Kong’s, fairness has traditionally been seen as treating everyone equally, giving them the same task with the same input under the same conditions for the same length of time. Not surprisingly, this is also the standard interpretation in measurement theory. In contrast, one of the underlying principles of the SBA initiative was that students should be assessed against criteria, rather than against other students (a core element of a standards-based as opposed to norm-referenced system). Teachers were also given the freedom to tailor the assessment task—its focus, timing, and grouping arrangement—to suit

---

7I am making a subtle but critical distinction here between the physical location of the assessment and the assessors/candidates (i.e., in the classroom, rather than an examination hall) and the philosophical orientation of the assessment. Simply changing the physical parameters of the assessment is not in itself sufficient to change the nature of the assessment. Many teachers reported they cannot “assess,” only “mark.” They feel unable to make a difference in teaching and learning, to respond to individual needs, because of community expectations of convergence and commonality. Teachers feel their assessment processes are expected to change, without the fundamental purposes being explicitly challenged. Such role conflict results in increasing stress and a decline in perceived teacher expertise.
their students’ very different language levels, interests, and needs. Congruent with the principles of assessment for learning, it was assumed that “fairness” meant every student should be given the opportunity to demonstrate his or her best (similar to Swain’s 1984, notion of “bias for best”\(^8\)). In this sense, fairness sometimes required students to be treated *unequally*, for example, students with very low levels of English might be asked to read a simple story, then assessed by a simple recount with perhaps extensive scaffolding to ensure they gain a sense of achievement and learn from the assessment process. It would be considered “unfair” to give the same task or support to very proficient students, as their language would not be extended, and they would invariably underperform. As Hamp-Lyons (2001) notes, this view of fairness is actually closer to the more traditional notion of “fairness” embodied in the classical examinations for the Chinese civil service, that is, that conditions should be consciously created to make opportunities open to all.

However, many teachers in the SBA trial, despite eagerly endorsing the rhetoric of assessment for learning, in practice found it extremely difficult to free themselves and their students from their existing conceptions. A significant minority of teachers were found to be treating SBA as if it were a separate, one-off externally set and assessed exam, albeit located in the school. Some were very concerned that “students will be using texts of different lengths and levels of difficulty. The assessment tasks may also be different from school to school. It doesn’t seem fair.”

Some teachers in the SBA study perceived the differences between the old and the new assessment cultures as a major stumbling block to assessment reform, exemplified by the following comment:

> I feel it takes time because the culture, the education culture in Hong Kong is different from other western countries and the students may not be used to that kind of assessment. They like to do exam paper. They think (then) they have something to learn.

However, another teacher newly returned from overseas, commented:

> Actually I was so surprised that … how slow we Hong Kong people are in terms of education … because I remembered when I was in Canada, we never … you would never be graded on just one exam. It’s quite like what

---

\(^8\)Swain (1984) advocated that to elicit the best performance from students in a communicative classroom, the following conditions must be met: more than adequate time to complete a task, opportunity given to review and change work, access given to reference materials, checking that students are on task, clear instructions including what is being assessed, and useful suggestions about how to do the task.
we are trying to do actually. I believe that (assessment for learning) has been practised in those places for years and I was actually surprised nobody did anything (here). … I am totally for assessment for learning.

Not surprisingly, the trial schools that were already doing extensive reading and whose students engaged in oral group work and individual presentations on a regular basis found it the easiest to integrate the assessment tasks into their existing practice, albeit with adaptations to meet the SBA guidelines. For example, one teacher commented:

I just briefly tell the students about the task because it is in mid May, so they were quite busy that moment. So I asked them just make use of they been doing say they just, they can just took from ERS report and work on it, say prepare a better review so that they can just have their presentation based on the review and I told them the date and the time of presentation. That’s what I did at the very beginning. Later on, I, I met them some days later, and I asked them to show me the book review they had written and I took a very look at it and I found that there weren’t any major problems in it. So I just returned them the review and they started to prepare those tasks … and later on, just right before they did the presentation, I helped them with the vocabulary and the names because they didn’t know how to pronounce them. So I just helped them pronounce them correctly… (but) I gave them more guidance according to the SBA documents … because the five questions listed there suggested some sort of high order thinking skills. … So I try to scaffold them to think in that way.

However, the trial showed that for many teachers, implementing SBA involved a steep learning curve and a significant change in their approach to teaching and to their role in student assessment. Responses were quite varied between schools and sometimes between teachers in the same school. Some teachers seemed very enthusiastic right from the beginning and stretched their creativity in thinking about how to build on the opportunities offered by SBA. A few teachers took much longer to come to grips with the principles involved in SBA, and their implications for teaching and learning as well as for assessment practice, as can be seen from the following quote from a trial school:

For students of higher forms, the time (eight minutes) is quite limited. They can’t have enough time to introduce their books and ask each other questions.

In fact, the SBA guidelines asked teachers to set their own time limits according to the needs of the students, but this was interpreted through the prism of
teachers’ existing experience—many schools used buzzers and stopwatches to allocate an identical period of time to each student, with the result that in some schools students’ stress levels were high and their “performance” very contrived and/or rushed.

As an outcome-oriented standards-referenced system, SBA is a significant cultural and attitudinal change, not only for teachers but for the whole school community, including students and parents. Hence, it is not surprising that fairness was a deep-seated sociocultural, not just political concern.

Technical Issues

At the technical level, as in Brindley (1998), concerns with SBA expressed by teachers and students revolved around the understanding and interpretation of traditional concepts such as reliability, validity, and authenticity. As indicated earlier, SBA is by its very nature teacher mediated, co-constructed and dialogic, context-dependent, multiple and varied, and dynamic and evolving (Black & Wiliam, 1998; Brookhart, 2003; Stiggins, 2001). In many ways SBA is the opposite to traditional testing in which context is regarded as an extraneous variable that must be controlled and neutralized and the assessor seen as someone who must remain objective and uninvolved throughout the whole assessment process (Lynch, 2001). SBA, in contrast, derives its validity from its location in the actual classroom where assessment activities are embedded in the regular curriculum and assessed by a teacher who is familiar with the student’s work. Like qualitative research, SBA builds into its actual design the capacity for triangulation, the collection of multiple sources and types of evidence under naturalistic conditions over a lengthy period of time. Hence, Rea-Dickins (2006) argues:

The traditional positivist position on language testing, with the tendency to map the standard psychometric criteria of reliability and validity onto the classroom assessment procedures, has been called into question, and the scope of validity has been significantly broadened (e.g., Chapelle, 1999; Lynch, 2001, 2003; McNamara, 2001) and taken further by a number of researchers. (p. 512)

Teasdale and Leung (2000), attempting to clarify the epistemological bases of different types of assessment in relation to spoken English language assessment in mainstream multiethnic classrooms in England, observed that alternative and psychometrically oriented assessment derive from different intellectual sources. Extending this argument, Leung (2004a, 2004b) argues that the evaluation criteria traditionally associated with psychometric testing such as reliability and validity need to be reinterpreted in SBA. Without such reinterpretation, SBA may be reduced to the testing of linguistic knowledge through a series of summative mini-achievement tests encouraging rehearsed dialogues with little or no opportunity
for spontaneous language use and “disembedded from the flow of teaching and learning” (Rea-Dickins, 2006, p. x).

On the other hand, some language testing researchers would argue that traditional test criteria do apply to alternative assessment schemes such as the SBA; for example, Clapham (2000) observed

A problem with methods of alternative assessment, however, lies with their validity and reliability: Tasks are often not tried out to see whether they produce the desired linguistic information; marking criteria are not investigated to see whether they ‘work’; and raters are often not trained to give consistent marks. (p. 152)

These same debates over technical issues arose, and continue to arise, with the HKEAA SBA initiative, exemplified by some teachers’ concerns that students would “cheat”—by memorizing whole chunks of spoken text, by overrehearsing and/or requesting they redo the same task again and again, or by blindly parroting their partners’ responses in group interaction. Teachers were also concerned that other teachers would “cheat” by allowing students to take copious notes or read aloud, or by giving them the same task again and again, or by simply fabricating results. Favouritism or bias was also seen as an issue by both teachers and students. More common was a lack of confidence among teachers, with many doubting that they had the required knowledge and skills to carry out the assessment properly. As one teacher commented, “I would like the HKEAA to take up my marks to see if I have interpreted the criteria correctly.” The notions of peer debriefing and standardisation were foreign to many teachers, and they saw such processes as just adding to their workload.

However, despite many concerns about reliability and validity, the teachers who were involved in the trial of the SBA did observe that they were surprised at the naturalness and ease with which many of their students were able to communicate. They commented that as the SBA was based on independent and extensive reading/viewing, students often neither had all read or viewed the same material nor had the same ideas. Thus when individuals made presentations or interacted with a group, there was often a genuine information gap, hence real interest among the students in what others had to say, enhanced by the familiar, more relaxing surroundings of the classroom. These teacher perceptions were reinforced by similar comments by the students themselves.

Another factor identified by the informants as contributing to greater authenticity in speaking was the actual assessment criteria and processes. For example, initially some teachers assumed that to do well on the SBA tasks it was necessary to rehearse students as if for an external exam (and an external examiner). However, the trial showed very clearly that students who relied on extensive notes or memorization did much less well in terms of communication strategies. Their pronunciation and language also suffered when they tried to memorize material
they had not fully mastered, thus leading to a tendency to mispronounce words, produce unnatural intonation, and lose their train of thought. As the assessment was undertaken by the class teacher, who was familiar with the range of student performance and who could ask questions to ensure the text was a student’s own work, many teachers eventually concluded that there was little possibility of cheating.

The teachers increasingly felt that the reliability of the assessment was also enhanced by having a series of assessments (rather than just one) by a teacher who was familiar with the student and by encouraging multiple opportunities for assessor reflection and standardisation. The exemplars of student work (a CD-ROM set was developed for use as a training package for teacher scoring and standardisation) provided a starting point for discussing the set standards. However, even after the trial some teachers assumed that achieving superficial consensus was the key to reliability rather than understanding and interpreting the assessment criteria and being able to justify a score to others. As a number of educational researchers have argued, developing the capacity and confidence to disagree can actually create more reliable conditions for assessment, because it allows misunderstandings and inconsistencies in the interpretation of the criteria to emerge and be challenged in a familiar and supportive environment (Davison, 2004; Moss, 1994). However, in the trial it was clear that Hong Kong teachers would not naturally adopt such a position; rather, they needed explicit encouragement and modelling, suggesting this was as much a sociocultural issue as a technical one.

Practical Issues

It is perhaps not surprising, given what is known from the literature on educational innovation and change (e.g., Cheung, 2001; Cheung & Ng, 2000), that it was the immediate and practical issues that most concerned the majority of teachers involved in the initial implementation of the SBA. Practical issues and concerns raised by teachers in relation to implementation of SBA included the following:

- The need for access to appropriate assessment (and extensive reading) resources.
- The need for activities and techniques as models/resources.
- Concerns about the type of recordings of oral performance that they were expected to collect.
- Lack of practical support for teachers at the school level.
- Concerns about the adequacy of professional development in SBA.
- Lack of time to implement and discuss assessments.
- Competing demands and priorities in relation to time allocation.
A particular source of misunderstanding early on was the technical requirements of SBA. Many teachers in the trial and subsequent focus group interviews had concerns about the level of technical resources and expertise that they thought were required. Some teachers reported that they did not have access to video cameras within their school. The SBA guidelines have been modified to emphasise even further that recording is optional, and that “homemade” audio-recording is sufficient for feedback and standardisation purposes. However, many participating teachers felt video-recording was necessary for them to be able to review students’ performances and make good judgements, especially when scoring group interaction with more than two students at one time.

Apart from the issues just mentioned, teachers were also concerned about whether they would be adequately prepared to implement SBA. Although training sessions were conducted on the overall design of the SBA component, and on the assessment criteria and standards, teachers also expressed a desire to better understand the underlying assumptions of SBA, especially how to integrate assessment into teaching and learning and how to set up effective assessment tasks. As a teacher associated with HKEdCity, one of the dedicated Web sites for SBA teacher support, accurately noted, “many teachers have an urgent need to view others’ practices and share experiences. … We can film the good lessons for teachers and analyse the lessons. We (need to) focus on teaching instead of assessment only.” In light of this request, more professional development was scheduled for all teachers teaching F4 English in 2005–06 so that they could be supported during the whole of the assessment period in the second semester. An SBA handbook and an introductory DVD and booklet were also produced (see http://web.hku.hk/~sbapro/), so teachers could become more familiar with the assessment process. In addition the system of district- and systemwide support was extended, with more training provided for the group coordinators of each cluster of schools. In general, teachers were also concerned that there was not sufficient recognition at the school level by principals and panel chairs of the time needed to discuss ideas and standardise assessments. With the official launch of teacher-support material along with the ongoing training sessions, it was hoped that awareness of the importance of SBA would be raised and more support would be given at the school level.

However, the teachers’ major concern with SBA was their perception that it was the last of too many new initiatives that they had to juggle, along with their busy schedules and heavy workload. Teachers were not placated by reassurances from the HKEAA that once SBA became a routine part of classroom activities, there would be no significant increase in the workload for students and teachers.

---

9The other is the author’s Faculty of Education Web site for SBA projects: http://web.hku.hk/~sbapro/.
Obviously much more input is needed to show teachers how to integrate assessment, teaching, and learning in schools with still very large classes and insufficient resources, but their concerns about workload and about the reform being implemented too quickly without adequate support were also very real (for a summary of media concerns at the time of the introduction of the SBA, see http://web.hku.hk/~sbapro/doc/Chinese%20news%20related%20to%20SBA.doc). In fact, these concerns over workload and time threatened to derail the whole SBA initiative until the HKEAA was finally given permission to adjust the timeline and provide more training and resources.

A POSITIVE IMPACT ON TEACHERS, LEARNERS AND SCHOOLS?

Despite these problems and issues with the design and implementation of SBA, the responses of teachers and students to the underlying philosophy of SBA and its emphasis on improving the quality of teaching and learning were generally very positive. A comment from a teacher involved in the trial was typical:

Personally, I enjoyed this trialling experience. I learnt how to judge the students through this activity. Moreover, my students tried to do the presentation based on the guiding questions given to them. Students found this presentation quite interesting and motivating. They learnt how to speak confidently and bravely during this assessment activity. They found this presentation rewarding since they can learn not only from the book but also through their actual participating experience.

Those who were involved in the trialling indicated that they were now much more aware of the values underlying SBA, the principles that they needed to work with to ensure their students gained the most from it, and the various options now available to them. For example, one teacher in a trial school commented:

After my students … had finished the presentation, I give them feedback and try to improve their performance and … I was really amazed by the response. … I thought that one of my students actually is not very good in English, but after the presentation she tried very hard to do it again and again. Feedback really works and I found that my students have improved a lot.

Other trial school teachers commented on the increase in confidence among students and the positive effect of the assessment activity on other language skills:
M: I think my student Sandy, just the one, just videotaped. She has great improvement … She thinks that is useful, very useful and she told me that … it also helps her reading. I think the student can really improve a lot after this trialling.

C: Well, my student has shown a great improvement in terms of confidence and English proficiency. They like talking to each other in English and they are not being afraid of being videotaped.

A: Yeah, they start, they start reading English books as well.

S: I think is just like what you said, is very good training for confidence and the students actually articulate what they read and … they found reading very useful, purposeful, that is something that you can share with someone, is not something just happening in your inner self. So I think the one that I trained out of the two, one girl called, actually she is a very shy girl after that I found she gained some confidence and she does quite well even in her writing. So she did well in this writing exam. I’m not sure if it is the effect of that training experience (but) it seems that she gained some confidence during this period of time.

Students also commented very favourably on the assessment activities, as the following extract from an interview with students illustrates:

T: What do you think about the assessment task you did in presentation task 5?

S1: It was quite interesting that we need to think about what the character needs. We can buy a gift.

S2: I just think it’s easy to handle it.

T: Why?

S2: It’s quite interesting to think for a gift to the character.

T: When you are thinking of a gift is it difficult?

S2: I don’t think so because I can think of many, many gifts to solve the problem.

T: Did you enjoy working with your partners, why?

S2: Yes, I did because my partners are all my best friends. We didn’t have any gaps so we did the project perfect.

T: How about you?

S3: I also very enjoy doing the task with my friends as they know me very much. When I don’t know what can I say they will help me to continue the conversation.

S4: With the partner I won’t feel nervous.

Overall, one of the more significant benefits of the SBA initiative identified by students and teachers was the capacity for the students to comment on their
own development and to receive constructive feedback immediately after the assessment had finished, hence improving learning. Teachers commented that when examinations are externally set and assessed, the only feedback that students and teachers receive is a grade at the end of the year, with no opportunities for interaction with the assessor and no chance to discuss how to improve. Teachers and students also commented that ongoing assessment encouraged students to work consistently. At the same time, it was seen as a source of concrete data for the evaluation of teaching and assessment practices. In contrast, examinations were seen as purely summative in purpose, leading to a focus on exam technique, rather than outcomes, and outside teachers’ control or even understanding. However, with SBA many teachers and students saw themselves as becoming partners in the assessment process. Although SBA involved only oral language assessment, some teachers reported that the SBA initiative had increased their level of awareness and skill in assessment planning, which they found readily transferable to other areas of the English language curriculum, and beyond.

CONCLUSIONS

The HKEAA SBA initiative is a major assessment reform that entails substantial change in school culture and structures as well as in pedagogic expectations among students, teachers, administrators, parents, and the wider community. It requires the development of content- and context-appropriate assessment activities, instruments, and procedures that are explicitly linked to high-quality teaching and learning and English language teachers who are not only confident and skilled at making highly contextualized, consistent, and trustworthy assessment decisions but also effective at involving students in the assessment process. These are major challenges on both a theoretical and practical level, so it not surprising that this early research study at the very beginning of the initiative has been followed by a number of other studies looking at how teachers (and students) deal with specific aspects of the SBA initiative, including designing the assessment task, making grouping decisions, intervening in the assessment process, involving students in self- and peer assessment, giving feedback, making assessment judgments, and using summative assessments for formative purposes. There are also studies evaluating the impact of the SBA initiative on students, on teachers, and on the English language curriculum. The results of this extensive research will be the ultimate test of whether teachers and students’ initial concerns are a natural response to a new initiative or an indication of more deep-seated problems with the assessment reform itself.
ACKNOWLEDGMENTS

An earlier version of this article was presented at the 8th Asian Forum of English Language Testing Associations, Hong Kong, November 2005. I acknowledge the support of the HKEAA and the Hong Kong Research Grants Council (RGC HKU 7268/04H) for funding this research, and the extensive assistance with the study of the SBA Consultancy Team. I am particularly grateful for the constant advice and encouragement of my colleague and coinvestigator, Professor Liz Hamp-Lyons.

REFERENCES


Hamp-Lyons, L. (1999). Implications of the “examination culture” for (English language) education in Hong Kong. In V. Crew, V. Berry, & J. Hung (Eds.), Exploring diversity in the language curriculum (pp. 133–140). Hong Kong: Hong Kong Institute of Education.


APPENDIX A
Sample Assessment Tasks

This is one of the sample assessment tasks developed for the school-based assessment (SBA). Please visit the SBA Web site at http://web.hku.hk/~sbapro/sba.html for more sample assessment tasks.

<table>
<thead>
<tr>
<th>Name of Task: New Neighbours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral Text-type:</td>
</tr>
</tbody>
</table>

**Communication Functions:**
- ☑ describing
- ☑ reporting
- ☑ explaining
- ☐ discussing
- ☐ classifying
- ☐ comparing
- ☐ persuading
- ☐ others:

**Audience--teacher plus:**
- ☐ a student partner
- ☐ small groups
- ☑ class
- ☐ more than one class

**Targeted audience:**
- ☑ fellow students
- ☐ students from other classes
- ☐ teacher(s)
- ☐ others: _______________

**Role(s) of audience:**
- ☐ giving non-verbal responses only
- ☑ questioning/commenting
- ☐ interacting with no limitations

**Where on this continuum would you place the task?**

- ☑ spontaneous, informal talk, e.g. small group interaction
- ☐ interactive, planned yet dialogic, e.g. report, interactive factual report
- ☑ individual long turn of planned, spoken text, e.g. news reporting, story telling
- ☐ individual long turn that is planned, cohesive, organized, formal, e.g. spoken report, a speech

**This task is suitable for use with the following genre(s):**
- ☑ print/non-print fiction
- ☑ print/non-print biography/autobiography
- ☐ factual books/documentaries on common topics, e.g. sports, hobbies, travel
- ☐ books/films on real life issues, e.g. environmental, social, economic

**Preparation:** none

**Description of pre-assessment activities:**
1. Ask students to think of an interesting character from a story/class reader that you have taught recently.
2. Ask them to imagine that one of the characters in the story has moved in next door to them.
3. Ask them to think about what life is like with such a neighbour.
4. Hold a discussion with the students and write down what kind of information they should cover if they were asked to describe an imaginary day they spent with the new neighbour. The information may include one or more of the followings:
   a) Name and gender of the neighbour
   b) How does he/she look like?
   c) How does he/she dress at home?
   d) What is his/her personality?
   e) How does he/she treat his family or people around him/her?
   f) What is/are the major event(s) in the story that your character takes part in?
   g) Do you like this new neighbour? Why/ Why not?
   h) How did you spend your day with this new neighbour? What did you do?
   i) What did you learn from this new neighbour?
5. For homework, ask each student to write a description of an imaginary day he/she spent with "the new neighbour".

---

This is a sample of the sample assessment tasks developed for the school-based assessment (SBA). It is recommended to visit the SBA Web site at http://web.hku.hk/~sbapro/sba.html for more sample assessment tasks.
6. Remind them to draw references from the books. They can’t turn their new neighbour into a wonderful person, if the descriptions from the book prove otherwise.
7. In the next lesson, ask students to share what they wrote in small groups.
8. Ask students to nominate the most interesting presentation among their group members.
9. Invite a student from each group to share their presentation with the whole class.

| Planned SBA Task: |
| Ask the students to describe an imaginary day in their lives when they spend time with a character from a book or film they viewed. Ask them to provide some background information about the book/film they read/viewed before they describe their imaginary day with their new neighbour during the individual oral presentation. |
| Tips/comments: |
| - The personal responses for this task can provide a good basis for discussion in English at a comfortable level. |
| - If students need more opportunities to speak in public, you may invite each student to take turns sharing their presentation with the whole class. |

| Sources: |
## APPENDIX B

### TABLE B1
School-Based Assessment Criteria for Group Interaction

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>6</strong> Can project the voice appropriately for the context. Can pronounce all sounds/sound clusters and words clearly and accurately. Can speak fluently and naturally, with very little hesitation, and using intonation to enhance communication.</td>
<td>Can use <strong>appropriate</strong> body language to <strong>display and encourage</strong> interest. Can use a <strong>full range</strong> of turn-taking strategies to initiate and maintain appropriate interaction and can <strong>draw others</strong> into extending the interaction (e.g., by summarising for others’ benefit, or by redirecting a conversation); can <strong>avoid the use of narrowly formulaic expressions</strong> when doing this.</td>
<td>Can use a <strong>wide range of accurate vocabulary</strong>. Can use varied and highly accurate language patterns; minor slips do not impede communication. Can <strong>self-correct effectively</strong>.</td>
<td>Can express a <strong>wide range</strong> of relevant information and ideas without any signs of difficulty. Can consistently <strong>respond effectively</strong> to others, sustaining and extending a conversational exchange. Can use the <strong>full range of questioning and response levels</strong> (see Framework of Guiding Questions) to engage with peers.</td>
</tr>
<tr>
<td><strong>5</strong> Can project the voice appropriately for the context. Can pronounce all sounds/sound clusters clearly and almost all words accurately.</td>
<td>Can use <strong>appropriate</strong> body language to <strong>display and encourage</strong> interest. Can use a <strong>good range</strong> of turn-taking strategies to <strong>initiate and maintain appropriate interaction</strong> (e.g., by encouraging contributions from others’ in a group discussion, by asking for others’ opinions, or by responding to questions); can <strong>mostly avoid the use of narrowly formulaic expressions</strong> when doing this.</td>
<td>Can use varied and almost always appropriate vocabulary. Can use almost entirely accurate and appropriate language patterns.</td>
<td>Can express relevant information and ideas clearly and fluently. Can respond appropriately to others to <strong>sustain and extend</strong> a conversational exchange.</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Can</strong> speak fluently with only occasional hesitation, and using intonation to enhance communication, giving an overall sense of natural nonnative language.</td>
<td><strong>Can</strong> use <strong>some features of appropriate body language</strong> to encourage and display interest.</td>
<td><strong>Can</strong> use <strong>some creative as well as formulaic</strong> expressions if fully engaged in interaction.</td>
<td><strong>Can use</strong> a <strong>good variety of questioning and response levels</strong> (see Framework of Guiding Questions).</td>
</tr>
<tr>
<td><strong>Can</strong> project the voice mostly satisfactorily.</td>
<td><strong>Can use</strong> a <strong>range of appropriate turn-taking strategies</strong> to participate in, and sometimes initiate, interaction (e.g., by responding appropriately to others’ comments on a presentation, by making suggestions in a group discussion).</td>
<td><strong>Can</strong> use <strong>some language patterns</strong> that are usually accurate and without errors that impede communication.</td>
<td><strong>Can present</strong> relevant literal ideas clearly with well-organised structure.</td>
</tr>
<tr>
<td><strong>Can</strong> pronounce most sounds/sound clusters and all common words clearly and accurately; less common words can be understood although there may be articulation errors (e.g., dropping final consonant clusters).</td>
<td><strong>Can</strong> use mostly appropriate vocabulary.</td>
<td><strong>Can</strong> self-correct when concentrating carefully or when asked to do so.</td>
<td><strong>Can often respond appropriately</strong> to others; <strong>can sustain and may extend</strong> some conversational exchanges.</td>
</tr>
<tr>
<td><strong>Can</strong> speak at a deliberate pace, with some hesitation but using sufficient intonation conventions to convey meaning.</td>
<td><strong>Can</strong> usually self-correct effectively.</td>
<td><strong>However:</strong> <strong>Can</strong> do these things less well when attempting to respond to interpretive or critical questions, or <strong>can interpret</strong> information and <strong>present elaborated ideas</strong>, but at these questioning levels coherence is not always fully controlled.</td>
<td><strong>However:</strong> <strong>Can</strong> do these things less well when attempting to respond to interpretive or critical questions, or <strong>can interpret</strong> information and <strong>present elaborated ideas</strong>, but at these questioning levels coherence is not always fully controlled.</td>
</tr>
</tbody>
</table>
Volume may be a problem.
Can pronounce all simple sounds clearly but some errors of sound clusters; less common words may be misunderstood unless supported by contextual meaning.
Can speak at a careful pace and use sufficient basic intonation conventions to be understood by a familiar and supportive listener; hesitation is present.

Can use appropriate body language to show attention to the interaction.
Can use appropriate but simple and formulaic turn-taking strategies to participate in, and occasionally initiate, interaction (e.g., by requesting repetition and clarification or by offering praise).
Can use simple vocabulary and language patterns appropriately and without errors that impede communication.
Can sometimes self-correct simple errors.
May suggest a level of proficiency above 3 but has provided too limited a sample.
Can present some relevant ideas sequentially with some links among their own ideas and with those presented by others.
Can respond to some simple questions and may be able to expand these responses when addressed directly.

Volume may be a problem.
Can pronounce simple sounds/sound clusters well enough to be understood most of the time; common words can usually be understood within overall context.
Can produce familiar stretches of language with sufficiently appropriate pacing and intonation to help listener’s understanding.

Can use appropriate body language when especially interested in the group discussion or when prompted to respond.
Can use simple but heavily formulaic expressions to respond to others (e.g., by offering greetings or apologies).
Can use some very basic language patterns accurately in brief exchanges.
Can identify some errors but may be unable to self-correct.
Provides a limited language sample.
Can express some simple relevant information and ideas, sometimes successfully, and may expand some responses briefly.
Can make some contribution to a conversation when prompted.

Volume is likely to be a problem.
Can pronounce some simple sounds and common words accurately enough to be understood.
Can use appropriate intonation in the most familiar of words and phrases; hesitant speech makes the listener’s task difficult.

Can use restricted features of body language when required to respond to peers.
Can use only simple and narrowly restricted formulaic expressions, and only to respond to others.
Can produce a narrow range of simple vocabulary.
Can use a narrow range of language patterns in very short and rehearsed utterances.
A restricted sample of language makes full assessment of proficiency difficult.
Can occasionally produce brief information and ideas relevant to the topic.
Can make some brief responses or statements when prompted.

Does not produce any comprehensible English speech.
Does not use any interactional strategies.
Does not produce any recognisable words or sequences.
Does not produce any appropriate, relevant material.
|------------------------------------------------|----------------------------|----------------------------------|-------------------------|
| 6  **Can** project the voice appropriately for the context.  
**Can** pronounce all sounds/sound clusters and words clearly and accurately.  
**Can** speak fluently and naturally, with very little hesitation, and using intonation to enhance communication. |  **Can** use appropriate body language to show focus on audience and to engage interest.  
**Can** judge timing to complete the presentation.  
**Can** confidently invite and respond to questions or comments when required for the task. |  **Can** use a wide range of accurate vocabulary.  
**Can** use varied and highly accurate language patterns; minor slips do not impede communication.  
**Can** choose appropriate content and level of language to enable audience to follow, without the use of notes.  
**Can** self-correct effectively. |  **Can** convey relevant information and ideas clearly and fluently without the use of notes.  
**Can** elaborate in detail on some appropriate aspects of the topic, and can consistently link main points with support and development. |
| 5  **Can** project the voice appropriately for the context.  
**Can** pronounce all sounds/sound clusters clearly and almost all words accurately.  
**Can** speak fluently with only occasional hesitation, and using intonation to enhance communication, giving an overall sense of natural nonnative language. |  **Can** use appropriate body language to show focus on audience and to engage interest.  
**Can** judge timing sufficiently to cover all essential points of the topic.  
**Can** appropriately invite and respond to questions or comments when required for the task. |  **Can** use varied and almost always appropriate vocabulary.  
**Can** use almost entirely accurate and appropriate language patterns.  
**Can** choose content and level of language that the audience can follow, with little or no dependence on notes.  
**Can** usually self-correct effectively. |  **Can** convey relevant information and ideas clearly and well.  
**Can** elaborate on some appropriate aspects of the topic, and can link main points with support and development. |
4 Can project the voice mostly satisfactorily. Can pronounce most sounds/sound clusters and all common words clearly and accurately; less common words can be understood although there may be phonation errors (e.g., dropping final consonant clusters). Can speak at a deliberate pace, with some hesitation but using sufficient intonation conventions to convey meaning. Can use appropriate body language to display audience awareness and to engage interest, but this is not consistently demonstrated. Can use the available time to adequately cover all the most essential points of the topic. Can respond to any well-formulated questions that arise. Can use mostly appropriate vocabulary. Can use language patterns that are usually accurate and without errors that impede communication. Can choose mostly appropriate content and level of language to enable audience to follow, using notes in a way that is not intrusive. Can self-correct when concentrating carefully or when asked to do so. Can present relevant literal ideas clearly and in well-organised structure. Can expand on some appropriate aspects of the topic with additional detail or explanation, and can sometimes link these main points and expansions together effectively.

3 Volume may be a problem. Can pronounce all simple sounds clearly but some errors of sound clusters; less common words may be misunderstood unless supported by contextual meaning. Can speak at a careful pace and use sufficient basic intonation conventions to be understood by a familiar and supportive listener; hesitation is present. Can use some appropriate body language, displaying occasional audience awareness and providing some degree of interest. Can present basic relevant points but has difficulty sustaining a presentation mode. Can respond to any cognitively simple, well-formulated questions that arise. Can use simple vocabulary and language patterns appropriately and without errors that impede communication, but reliance on memorised materials or written notes makes language and vocabulary use seem more like written text spoken aloud. Can choose a level of content and language that enables audience to follow a main point, but needs to refer to notes. Can sometimes self-correct simple errors. Can present some relevant literal ideas clearly, and can sometimes provide some simple supporting ideas. Can sometimes link main and supporting points together.

(Continued)
**TABLE B2**
(Continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2</strong></td>
<td>Can use a restricted range of features of body language, but the overall impression is stilted.</td>
<td>Can appropriately use vocabulary drawn from a limited and very familiar range.</td>
<td>Can make an attempt to express simple relevant information and ideas, sometimes successfully, and can attempt to expand on a few points.</td>
</tr>
<tr>
<td>Volume may be a problem.</td>
<td>Can present very basic points but does not demonstrate use of a presentation mode and is dependent on notes.</td>
<td>Can read notes aloud but with difficulty.</td>
<td>Can link the key information sequentially.</td>
</tr>
<tr>
<td>Can pronounce simple sounds/</td>
<td>Audience awareness is very limited.</td>
<td>Can use some very basic language patterns accurately in brief exchanges.</td>
<td></td>
</tr>
<tr>
<td>sound clusters well enough to be understood most of the time; common words can usually be understood within overall context.</td>
<td></td>
<td>Can identify some errors but may be unable to self-correct.</td>
<td></td>
</tr>
<tr>
<td>Can produce familiar stretches of language with sufficiently appropriate pacing and intonation to help the listener’s understanding.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>1</strong></td>
<td>Body language may be intermittently present, but communication strategies appropriate to delivering a presentation are absent. The delivery is wholly dependent on notes or a written text. There is no evident audience awareness.</td>
<td>Can produce a narrow range of simple vocabulary.</td>
<td>Can express a main point or make a brief statement when prompted, in a way that is partially understandable.</td>
</tr>
<tr>
<td>Volume is likely to be a problem.</td>
<td></td>
<td>Can use a narrow range of language patterns in very short and rehearsed utterances.</td>
<td></td>
</tr>
<tr>
<td>Can pronounce some simple sounds and common words accurately enough to be understood.</td>
<td>A restricted sample of language makes full assessment of proficiency difficult.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can use appropriate intonation in the most familiar of words and phrases; hesitant speech makes the listener’s task difficult.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>0</strong></td>
<td>Does not attempt a presentation.</td>
<td>Does not produce any recognisable words or sequences.</td>
<td>Does not express any relevant or understandable information.</td>
</tr>
<tr>
<td>Does not produce any comprehensible English speech.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## APPENDIX C

### Assessment Record

**TABLE C1**

**Hong Kong Certificate of Education Examination English Language School-Based Assessment Component: Assessment Record (Group Interaction)**

<table>
<thead>
<tr>
<th>School Name:</th>
<th>Oral Text-type: Group Interaction</th>
<th>Assessment date: <em><strong>/</strong></em>/___</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher Name:</td>
<td>Name of text:</td>
<td>Category: Print / N-Print</td>
</tr>
</tbody>
</table>

(circle)

Name of text: ________________________________

(category)

Assessment date: ___/___/___

Name of text: ________________________________

Category: Print / N-Print

(circle)

Assessment date: ___/___/___

Name of text: ________________________________

(circle)

Assessment date: ___/___/___

Name of text: ________________________________

(circle)

**ADVICE TO TEACHERS**

This assessment sheet will assist teachers to allocate marks. There are two stages to this process. The first stage is to make judgments on the student’s performance in each domain (i.e. pronunciation and delivery, communication strategies, vocabulary and language patterns, and ideas and organisation) with reference to the Assessment Criteria. You should circle one of the numbers 1–6 (or 0 if no language was produced) to indicate how well the student performed in each domain. The second stage is to add up the marks for all domains. The total number of possible marks is 24. Add a comment if possible.

<table>
<thead>
<tr>
<th>Student 1: No.</th>
<th>Student 2: No.</th>
<th>Student 3: No.</th>
<th>Student 4: No.</th>
</tr>
</thead>
</table>

**CRITERIA FOR THE AWARD OF MARKS**

(Circle number for each domain)

1. Pronunciation & delivery
   0 1 2 3 4 5 6

2. Communication strategies
   0 1 2 3 4 5 6

3. Vocabulary & language patterns
   0 1 2 3 4 5 6

4. Ideas & organisation
   0 1 2 3 4 5 6

**TOTAL: _____ / 24**

<table>
<thead>
<tr>
<th>Student 3: No.</th>
<th>Student 4: No.</th>
</tr>
</thead>
</table>

**CRITERIA FOR THE AWARD OF MARKS**

(Circle number for each domain)

1. Pronunciation & delivery
   0 1 2 3 4 5 6

2. Communication strategies
   0 1 2 3 4 5 6

3. Vocabulary & language patterns
   0 1 2 3 4 5 6

4. Ideas & organisation
   0 1 2 3 4 5 6

**TOTAL: _____ / 24**

**TEACHER’S COMMENTS**

**AUTHENTICATION**

1. I certify that each student has read/viewed the text above used in this oral assessment, that the text is not a class reader, comic, newspaper, or a set text for other subjects, and that the work is all the student’s own.

2. I certify that the assessment was undertaken under the conditions specified in the HKEAA guidelines, that I am the student’s English teacher, that I conducted the assessment and that the task has not been repeated.

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Student 1</th>
<th>Student 2</th>
<th>Student 3</th>
<th>Student 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signature</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**ADVICE TO TEACHERS**

This assessment sheet will assist teachers to allocate marks. There are two stages to this process. The first stage is to make judgments on the student’s performance in each domain (i.e. pronunciation and delivery, communication strategies, vocabulary and language patterns, and ideas and organisation) with reference to the Assessment Criteria. You should circle one of the numbers 1-6 (or 0 if no language was produced) to indicate how well the student performed in each domain. The second stage is to add up the marks for all domains. The total number of possible marks is 24. Add a comment if possible.

**CRITERIA FOR THE AWARD OF MARKS**

(Circle number for each domain)

<table>
<thead>
<tr>
<th>Domain</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Pronunciation &amp; delivery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Communication strategies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Vocabulary &amp; language patterns</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Ideas &amp; organisation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TOTAL:** __________ / 24

**TEACHER’S COMMENTS**

Comments on aspects of the student’s work that led to your assessment and any contextual factors (e.g. amount of rehearsal or teacher support) that need to be taken into account.

**AUTHENTICATION**

1. I certify that this student has read/viewed the text above used in this oral assessment, that the text is not a class reader, comic, newspaper, or a set text for other subjects, and that the work is all the student’s own.

2. I certify that the assessment was undertaken under the conditions specified in the HKEAA guidelines, that I am the student’s English teacher, that I conducted the assessment and that the task has not been repeated.

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signature</td>
<td>Signature</td>
</tr>
<tr>
<td>Date</td>
<td>Date</td>
</tr>
</tbody>
</table>