Chapter 6
The English Language Benchmark Subject Committee

David Coniam and Peter Falvey

Abstract This chapter details the work of the English Language Benchmark Subject Committee (ELBSC) in developing, moderating, amending, changing and overseeing further the work on language benchmark developments.

The English Language Benchmark Subject Committee—Purpose and Brief

The English Language Benchmark Subject Committee (ELBSC) was convened in October 1997 under the auspices of the then Hong Kong Examinations Authority (HKEA) which later became the Hong Kong Examinations and Assessment Authority (HKEAA). Its purpose was to produce language benchmark specifications and an assessment syllabus for promulgation to Hong Kong teachers of English language in preparation for a large-scale pilot exercise—the Pilot Benchmark Assessment (English) (PBAE). The objective of the PBAE was to examine the prototype benchmark tests which the ELBSC had recommended, and to trial these tests on as representative a sample as possible of the Hong Kong English language teacher cohort. The composition of the ELBSC was very broad. The time frame the ELBSC was given was one year, using the consultancy report (Coniam & Falvey, 1996) as the starting point for the ELBSC's initial discussions. There was considerable debate over the substance of the report. While the majority of the recommendations were accepted by the ELBSC—that is the areas to be assessed—certain details of how assessment might be accomplished—the format of the Reading and Listening Tests,

D. Coniam (✉) · P. Falvey
Department of Curriculum and Instruction, Faculty of Education & Human Development,
The Education University of Hong Kong, 10 Lo Ping Road, Tai Po, Hong Kong
e-mail: coniam@eduhk.hk

P. Falvey
e-mail: falvey@eduhk.hk

for example, and the scales and descriptors of the CLA, were not wholly accepted by the ELBSC. To resolve questions that the ELBSC raised, five Working Party sub-groups were formed under the ELBSC, each tasked with investigating one of the five areas to be assessed, namely Reading, Writing, Speaking, Listening and Classroom Language Assessment.

Pre-PBAE Validation Studies

For some of the test types, validation exercises of the test material, or of the training and standardisation of assessors were conducted by the consultants, with subjects consisting of in-service and pre-service teachers at local universities. Eight reports were produced by the consultants focusing on the validation of the assessment instruments and the training and standardisation of assessors for the criterion-referenced tests.

The reports contain detailed information on different aspects of the development of the English Language Benchmarking Initiative. These reports were:

1. Validating the Classroom Language Assessment Component: The Hong Kong English Language Benchmarking Initiative (Coniam & Falvey, 1998a)
2. Validating the Reading Test: The Hong Kong English Language Benchmarking Initiative (Coniam & Falvey, 1998b)
3. Piloting the Multiple-Choice Cloze Test: The Hong Kong English Language Benchmarking Initiative (Coniam & Falvey, 1998c)
4. Validating the Speaking Test: The Hong Kong English Language Benchmarking Initiative (Coniam & Falvey, 1998d)
5. Pre-pilot Exercise Rewriting and Speaking Components of the English Language Benchmark Project (Falvey & Coniam, 1998a)
6. Assessor Training and Standardisation for Classroom Language Assessment: The Hong Kong English Language Benchmarking Initiative (Falvey & Coniam, 1998b)
7. Assessor Training and Standardisation for the Speaking Test: The Hong Kong English Language Benchmarking Initiative (Falvey & Coniam, 1999c)
8. Assessor Training and Standardisation for the Writing Test: The Hong Kong English Language Benchmarking Initiative (Falvey & Coniam, 1999d).

The studies will be referred to from time to time in this and the following chapters.

As the list above reveals, no trialling of any material for the Listening Test was possible. The Listening Test that emerged suffered, not surprisingly, from the deficiencies that are discussed further in this chapter below.

The Work of the ELBSC

Between October and December 1997, the ELBSC met 32 times. The discussion and recommendations made by the ELBSC for test types are now described. A considerable number of amendments and changes—as might be expected—were made to the original recommendations of the 1996 consultancy feasibility study report by the ELBSC as a result of their deliberations.

Classroom Language Assessment

The CLA was discussed at length in the ELBSC because it would be a performance-based test that would take place in a live taught class. While the ELBSC was very much in agreement with the philosophy behind the use of an authentic test, logistic concerns were expressed at the administration of a live CLA.

Although English language teachers are used to paper-and-pencil tests, a live classroom test would be much more threatening. The constructs assessed would need to be broad in terms of language skills that were assessed, i.e. that they should not be biased against any particular group—primary versus secondary, for example. Care also had to be taken that the constructs which were to be established involved the assessment of language only and not pedagogical skills or personality traits. Support for the retention of CLA was made in a 1999 Colloquium on English Language Benchmarks held in Hong Kong, where Nevo (1999) stated unequivocally that the inclusion of the CLA in language benchmarking should be retained in spite of inevitable arguments that it would be costly and time-consuming.

A Working Party for CLA was formed under the main ELBSC to examine the constructs that the consultancy team had formulated in their original 1996 report and to examine the constructs, scales and descriptors both for validity and potential reliability. The Working Party met six times, watched over 20 videos, discussed the skills and constructs they felt appropriate to English language teachers, and reported back to the ELBSC.

There was strong agreement that the four constructs that had been formulated in the Consultancy Report for English Language Benchmarks (Coniam & Falvey, 1996) were the essential English language skills which teachers of English language required in order to underpin the effective teaching of English. *Grammatical Accuracy* and *Pronunciation, Stress and Intonation* are the two 'formal' elements which define an English language teacher's ability in English. The other two elements *The Language of Presentation/Practice* and *The Language of Interaction* are the functional realisations of a teacher's formal ability in English in terms of communicating with students and getting things done in the classroom. Scales and their descriptors were then formed to reflect those skills at various levels of ability.

The four constructs and their associated descriptors of language performance were arrived at by the following methods:

1. Observation of English language lessons on video,
2. Creation of a taxonomy of teacher language tasks,
3. Development of prototype constructs,
4. Moderation of the constructs by experts and practising teachers,
5. Creation of scales,
6. Creation of descriptors for each scale based upon distinct levels of language performance,
7. Validation of the constructs and descriptors through moderation and empirical study and
8. Submission of the prototypes to the ELBSC.

After phase (7), Level '3' of the prototype scales was adopted as the tentative benchmark level. A new Level '0' was added to indicate that no performance in that skill area was available for grading, e.g. speaking in Cantonese for the whole lesson. By mid-1998, the specifications of the scales after revision, modification and amendment were resolved as follows:

1. Grammatical Accuracy,
2. Pronunciation, Stress and Intonation,
3. The Language of Interaction and
4. The Language of Instruction.

The Speaking Test

Specifications

As reported in the validation study of the Speaking Test (Falvey & Coniam, 1999), the ELBSC agreed, after reviewing the different options proposed in the Consultants' Report for English Language Benchmarking (Coniam & Falvey, 1996), that the assessment of speaking was a crucial part of the English language benchmark assessment procedure.

In addition, the ELBSC eventually decided that some skills must be assessed for all teachers of English language, e.g. the comparatively difficult and teacher-specific skill of reading aloud; and the language teacher skill of storytelling or recounting. The 1996 Consultancy Report proposed three test types and seven separate scales for the Speaking Test.

Although the ELBSC'S deliberations on the Speaking Test retained the essence of the consultants' 1996 recommendations in that the test still consisted of three linked elements, certain elements and task types were changed.

As can be seen from Table 6.1, one of the original scales (Pronunciation, Stress and Intonation) was tested twice in the original proposals so one of those pronunciation scales was dropped. The test types and scales which the ELBSC accepted and on

Table 6.1   Scales and descriptors in the 1996 Consultancy Report

Test type	Scale	Salient linguistic features
1. Reading aloud: giving instructions	1. Pronunciation, Stress and Intonation	Sounds, stress, intonation
	2. Reading Aloud with Meaning	Speed of delivery, pausing, awareness of audience
2. Talking about teaching	1. Pronunciation, Stress and Intonation	Sounds, stress, intonation
	2. Grammatical Accuracy	Grammatical accuracy, range of structures
	3. Organisation and cohesion	Coherence, logical flow of ideas, relationships between ideas
3. Oral interaction	1. Interacting with Peers	Including turn-taking, initiating, responding, agreeing and disagreeing
	2. Explaining Language Matters to Peers	Including the use of appropriate metalanguage, appropriate examples

Table 6.2   Scales and descriptors proposed by the ELBSC

Test type	Scale	Salient linguistic features
1. Reading aloud a text	1. Pronunciation, Stress and Intonation	Sounds, stress, intonation
	2. Reading Aloud with Meaning	Speed of delivery, pausing, awareness of audience
2. Telling a story/recounting a personal experience/presenting arguments	1. Grammatical Accuracy	Grammatical accuracy, range of structures
	2. Organisation and Cohesion	Coherence, logical flow of ideas, relationships between ideas
3. Professional oral interaction	1. Interacting with Peers	Including turn-taking, initiating, responding, agreeing and disagreeing
	2. Explaining Language Matters to Peers	Including the use of appropriate metalanguage, appropriate examples

which the PBAE Speaking Test was based and which are reported here are presented in Table 6.2.

The six scales and the descriptors that were used in the PBAE are contained in Appendix E "Speaking Test Scales and Descriptors", p. 81–85 Chap. 5.

Assessor Training for the PBAE

An investigation was conducted into the reliability of the Speaking Test assessors.

The purpose of the training was to train and standardise assessors. However, this also involved conducting an analysis of the assessors' scores in order to deselect potentially weak assessors, those who might be too harsh or too lenient or those unable to apply the scales and descriptors consistently.

The 16 assessors were first given the marking schemes and rating scales and given time to read and digest them. They were then shown the first set of three applicants and asked to rate them without discussion. The purpose of this blind rating was to enable the consultants to observe how much initial variability there was amongst the assessors. Subsequently, this variability was to be compared with their performance on the final ratings carried out at the end of the day.

After the first, blind, rating session, the trainee assessors were then given detailed training for three more full sessions and standardisation feedback and follow-up after they had given their grades. While there was an initial wide range of marks, this was reduced through the day's training to a much narrower range. Assessor-to-model misfit was also substantially reduced. A detailed description of the training and standardisation procedure for the Speaking Test is provided in Falvey and Coniam (2000)

All assessors remarked that they felt the assessor training session had been remarkably well organised and that they had benefited in terms of being prepared for assessing teachers on the PBAE Speaking Test. Many of the assessors' suggestions were adopted and incorporated into the PBAE Speaking Test. Examples of changes include the addition of a poem to the prose passage in the Reading Aloud section.

The Writing Test

The ELBSC agreed that the original construct that had been formulated in the Consultancy Report for English Language Benchmarks (Coniam & Falvey, 1996) was an essential facet of the English language skills which teachers of English language require in order to underpin the effective teaching of English.

The ELBSC retained the Expository Writing test type reported in this chapter but increased the number of levels from five to six by making Level 0 a description of 'no performance upon which to make an assessment'. The ELBSC also introduced a new test type (Rewriting) with the result that the Writing Test that was used in the PBAE finally consisted of two test types and five scales (drawn from the original test type and the new test type).

The new test type was an innovative test of writing awareness and writing skill. It requires test takers to rewrite a student essay (typically a low-level Secondary 5 [Year 11] essay, written for the HKCEE examination). The purpose of the rewriting task was to demonstrate that test takers can not only understand the problems associated

Table 6.3 Scales and descriptors proposed by the ELBSC for the rewriting task

Test type	Scale	Salient linguistic features
1. Writing professionally	Grammatical Accuracy	Grammatical accuracy, range of structures
	Organisation and Coherence	Organisation of text, coherence
	Task Completion	All tasks requested in the stimulus must be completed
2. Rewriting a student composition	Vocabulary and Grammar	Grammatical accuracy, range of structures, appropriate lexical choice
	Organisation and Presentation of Facts/Information	Logical flow of ideas, relationships between ideas, retention of main facts/information from the original student text

with the writing of the composition but also that they have the requisite skills to allow them to rewrite it in an acceptable/exemplary manner.

The task was trialled and found to work well. Once preliminary descriptors had been established, a sample batch of rewrites was given to consultants and HKU language education specialists who, acting as assessors, were asked to read the rewritten text and then use the prototype descriptors to assign a benchmark level to each text. Although, at this stage, little training was given to assessors, the assessors reported favourably on their ability to operationalise the descriptors. Adjustments were made to the prototype descriptors based on feedback from these assessors during a pre-PBAE pilot in 1998. Concurrent validity for the rewriting task was high with the expository writing task ($r = .66$, $p < .001$). Concurrent validity was also high with the other skills (calibrated MC items—$r = .63$, $p < .001$); Speaking Test—($r = .89$, $p < .001$).

The scales and descriptors used in the PBAE Writing Test are shown in full in Appendix D "Writing Test Scales and Descriptors", p. 76–80. A summary is presented in Table 6.3.

The Reading Test

Reading

It was agreed, as for the Speaking Test, that a basic principle should be that teachers taking the test must be treated as mature adults and that multiple-choice tasks which resembled school tests should be avoided as far as possible. Principles laid down by the ELBSC were as follows:

1 It should tap higher-level reading skills.
2 It should neither duplicate HKEA school tests such as the HKCEE or HKASLE examinations nor appear similar to them. (This was for purposes of credibility and face validity. Teacher informants made it clear that they did not want to see a battery of tests which appeared to resemble the tests for which many of them were preparing their students.)
3 It should, ideally, *not* be in a multiple-choice format. (This criterion was established because of the ELBSC's desire to promote the more modern paradigm of assessment which eschews large-scale multiple-choice testing.)
4 The material should be authentic.
5 Its topic content should be based on domains that English language teachers might encounter in their professional lives, i.e. English language teaching and language education.

Cloze

The ELBSC also agreed eventually that a multiple-choice element should form part of the test battery.

Although initially resistant to the inclusion of multiple-choice test items, the ELBSC finally agreed to the inclusion of a multiple-choice cloze test because of the reliability such a test might afford the HKEA as an anchor against the Reading Test. The ELBSC stated, however, that:

1. The items should be integrated into a text type such as a cloze passage and not consist of discrete point items.
2. Some of the items should test discourse-level skills.
3. The items should be properly pretested.

Following the principle of using authentic material, a number of cloze passages were prepared for pretesting with as little amendment to their initial state as possible. Setters and moderators attempted to make as few amendments as possible to the original in order to provide teachers with the type of text that they could encounter in their professional lives. Item types included grammar and vocabulary, as do most cloze tests. However, an effort was also made to include items that required test takers to take the discourse context into account (c.f., Deyes, 1984).

The Listening Test

It was agreed that the stimulus for the Listening Test should consist of an authentic discussion, based around English language teaching/educational themes. It was decided not to use a single speaker as this would closely represent the academic listening skills required in a formal lecture. Consequently, a Listening Test was developed in which answers would be of an open-ended format. One of the ELBSC's recommendations was that the Listening Test should be delivered in a video rather than an audio format—the latter being the format adopted by the HKASLE Year 13 Use of English examination at the time.

A number of formats were experimented with. One of the formats involved the production of questions based on major themes rather than linear questions which paralleled the videotaped discussion. Participants in the video were briefed on the topic and then asked to take a stance on it. Topics covered included the use of native English-speaking teachers in Hong Kong secondary schools, the medium of instruction in schools and the role of English in education.

Unfortunately, the innovative video approach to a new test format was not piloted, due to logistical problems. As a result, the final format of the Listening Test can be described as a hybrid between a test for English language teachers and a 'more demanding' HKASLE Year 13 Use of English Listening Test. The majority of the questions generally paralleled the text (as with the Use of English Listening Test). Some questions did not, however, which required test takers to take a broader perspective, and to draw on different sections of the taped discussion. In addition, there was an attempt to include questions which required answers that drew on more than factual recall—the general item types used with questions in the HKASLE UE Listening Test.

As stated above, the wishes of the ELBSC were unable to be followed because of practical constraints (e.g. finding enough test rooms with video facilities for large numbers of test takers) so the HKEA decided that it would not be possible to administer the test which had been prepared for use on video via video. Instead, the video was converted to an audio tape for the live test. As reported below, this led to problems, reported by test takers. See also the discussion in Coniam (2001) of the relative lack of advantage of using video over purely audio as a medium for conducting listening tests and a further discussion of this topic in the closing chapter, Chap. 18.

Part 1: Assessment purpose / target group / objectives / language model
Part 2: Overall statement including a discussion of constructs
Part 3: Major components
Part 4: Task and question types
Part 5: Syllabus specifications (number of sections / papers / parts / suggested text lengths / timing etc.)

Fig. 6.1 Framework for Pilot Benchmark Assessment (English Language)

From the ELBSC and Its Working Parties to Its Moderation Committees

By the end of 1997, the ELBSC had agreed on the composition of the benchmark test battery, and a draft test blueprint was produced. This is now reproduced in Fig. 6.1.

Part 1: Assessment Purposes

Purpose of assessment	To establish minimum, acceptable levels of language ability for teachers of English in lower forms of secondary school
Target group	All teachers of English in lower forms of secondary school[a]
Objectives: Major	To establish minimum, acceptable levels of teacher language ability in order to underpin the effective teaching of English in lower secondary school classrooms
Objectives: Specific	To establish minimum, acceptable levels of competence in order to deliver the English language curriculum in the classroom in the specific language skill areas of: • Classroom language • Speaking • Listening • Reading • Writing
Language model	A functional model of language (Halliday, 1985) with reference for language testing to Bachman and Palmer's 1996 model of language (organisational [grammatical and textual] and pragmatic [functional and sociolinguistic]) knowledge and strategic (metacognitive strategies) competence

[a]The language skills of some upper primary and upper secondary teachers of English will also be sampled for purposes of comparison with their lower secondary counterparts

Part 2: Overall Statement Including a Discussion of Constructs

Construct Statement
The sections below contain construct descriptions of the major areas to be benchmarked. It will be noted that some overlap occurs. The reason for this is that some important language skills are used in different but relevant contexts of use. All of these contexts of use are deemed important for the effective practice of English both professionally (e.g., with colleagues and specialists) as well as in the classroom. Therefore, for example, it will be seen that the assessment of pronunciation occurs in two contexts—in a reading aloud task and in the context of the classroom with students. Grammar, likewise is assessed both in written (Writing component) and multiple-choice cloze form (Reading component) and, in addition, in two different but relevant spoken forms (speaking to peers/superiors and speaking to students in a live classroom context)

	Classroom Language Assessment
To be examined by	Education Department Classroom Language Assessors
Authenticity	
Areas to be benchmarked	CLASSROOM LANGUAGE in which minimum, acceptable levels of ability to communicate with students appropriately are assessed in the areas of: • Grammatical Accuracy • Pronunciation, Stress and Intonation • The Language of interaction • The Language of Instruction

	'Formal' Assessment: Direct and Indirect
To be examined by	Hong Kong Examinations Authority
Authenticity	In all cases, authentic texts, or adaptations of authentic material will be used
Areas to be benchmarked	SPEAKING in which competence is assessed, when interacting with educated native and non-native speakers, in the language skills areas of: • Pronunciation, Stress and Intonation • Reading Aloud with Meaning • Grammatical Accuracy • Organisation and Coherence • Interacting with Peers • Explaining Language Matters to Peers
	LISTENING in which competence is assessed by listening to and understanding educated native and non-native speakers of English in audio/video recordings. Possible text types would be discussions, debates, interviews, documentaries and current affairs programmes which discuss matters broadly related to education and professional language teaching. These might be drawn directly from the English language media in Hong Kong or developed from authentic interviews, discussions etc.
	READING in which competence in reading and understanding texts of an agreed appropriate nature and level within the context of professional language teaching is assessed (e.g., texts taken from journals such as Modern English Teacher, English Language Teaching Journal, Curriculum Forum, Practical English Teacher, as well as fiction and newspaper articles on relevant topics) VOCABULARY, GRAMMAR and DISCOURSE in which minimum, acceptable levels of vocabulary, grammar, discourse and textual knowledge are assessed in a cloze procedure
	WRITING in which competence is assessed by means of: • a stand-alone expository writing task • rewriting/improving a student composition

Part 3: Major Components

	Classroom Language Assessment
Areas to be assessed	Assessment of teacher language skills in a normal classroom working environment
Components/scales	**Classroom Language Assessment** To assess teachers' ability to use English for classroom purposes in the following ways: • Grammatically • With appropriate pronunciation, stress and intonation in order to demonstrate the communicative language skills which involve: • The Language of Interaction, i.e.: – Eliciting – Responding – Providing feedback – The language of classroom management, including: praising/advising/acknowledging • The Language of Instruction, i.e. – Presentation – Giving instructions – Signalling

	'Formal' Assessment: Direct and Indirect
Areas to be assessed	• Speaking • Listening • Reading, Vocabulary, Grammar and Discourse • Writing
Components/scales	SPEAKING • Pronunciation, stress and intonation • Reading aloud with meaning • Grammatical accuracy • Organisation and coherence • Interacting with peers • Explaining language matters to peers
	LISTENING Open-ended responses to audio/video-recorded spoken discourse
	READING Open-ended responses to texts VOCABULARY, GRAMMAR and DISCOURSE COMPONENT Multiple-choice cloze
	WRITING • Organisation and coherence • Grammatical accuracy • Task completion

Part 4: Task/Question Types

Classroom Language Assessment

A live lesson conducted with the teacher's normal time-tabled class which would include a display of the language skill areas which have been specified in previous sections

---

'Formal' Assessment: Direct and Indirect

SPEAKING
- An integrated "Reading and Speaking" task consisting of:
- – Reading aloud, e.g., narrative, instructions, poem etc., thematically linked to:
- – Telling a story/recounting a personal experience/presenting arguments based on a stimulus provided, e.g., written prompts, an incomplete short story, a set of pictures or the passage for reading aloud
- Discussing student language problems presented within the context of an authentic student composition

LISTENING
Open-ended listening tasks based on English language teaching situations or topics of general educational interest in the form of an exposition, e.g., lecture situation, dialogue or debate with the following types of question: factual; attitudinal; inferential; gist/summary

READING
- Open ended reading tasks based on a text or texts provided
VOCABULARY, GRAMMAR and DISCOURSE COMPONENT
- Multiple-choice cloze

WRITING
- An argumentative/explanatory/instructional writing task related to the professional or practical work of an English language teacher OR a writing task related to a text on a relevant language teaching topic
- Improving a student composition by identifying and solving lexico-grammatical and discourse problems

Part 5: Syllabus Specifications

Classroom Language Assessment

Paper IV: Classroom Language Assessment
Note: A minimum of 5 days' notice will normally be given by the assessor(s) to the teacher
Briefing: The teacher will brief the assessor(s) before the class takes place. The briefing will include information on the students' previous language learning and teacher language skills to be demonstrated
Time: 5–15 min, as required by the teacher
NOTE: This part is not assessed
Assessment: Live lesson:
The assessment will take place in a single period. The first 10 minutes of the single period will not be assessed. This non-assessed section of the lesson will allow the teacher, assessor(s) and students to get used to each other

	'Formal' Assessment: Direct and Indirect		
Major elements of the benchmarking assessment including: • number of sections • number of parts • text lengths • timing	PAPER-AND-PENCIL TESTS This assessment consists of two papers:		This assessment consists of two sections:
	Paper I Reading and Writing • Part 1: Multiple-choice cloze Time: 30 min Text: approximately 500 words Items: 20–30 • Part 2: Reading Time: 1 h Text(s): One text of 1500–2000 words or two texts of 750–1000 words each Questions: about 20 of various types • Part 3: Writing time: 1 h 15 min • Text: Stand-alone writing task: stimulus material will be given as input for the writing task, either using the reading passage in the reading comprehension component or a different text of 200–300 words. Text: Improving a student composition task—a text of about 200–300 words will be used	**Paper II Listening** • Listening and responding to an audio/video recording(s) which is/are heard only once Time: 1 h Preparation time: 3–5 min to look at the question paper Time for listening and responding: 30 min Completion time: 10–15 min • 'Text': One segment of spoken discourse of approximately 25–30 min or two segments of approximately 10–15 min each Questions: about 20 of various types	**Paper III Speaking** **Section A** consists of two parts: Preparation: 10 min Examination: 5 min • Reading aloud a text which will take a typical secondary school teacher of English one to two minutes to read aloud • Telling a story/recounting a personal experience/presenting arguments **Section B** consists of one part: Preparation: 5 min • Professional oral interaction Time: 10 min for discussion in group format Text: a student composition of about 200–300 words

As the HKSAR Government wished to press ahead with the specimen material and prepare for the PBAE (see below), in early 1998 four Moderation Committees were formed under the aegis of the Hong Kong Examinations Authority to set two sets of test material for the four paper-and-pencil tests. One set was to be released as specimen material to teachers; the other set was to be live pilot test material. A booklet of the prototype benchmark syllabus together with specimen material was published in September 1998 by ACTEQ in the syllabus document *Syllabus Specifications, Specimen Questions, Notes for Classroom Language Assessment*.

At the same time, i.e. September 1998, the Education Bureau began canvassing schools in an attempt to recruit teachers to participate in the Pilot Benchmark Assessment (English) , which is discussed in Chap. 7.

Summary

This chapter has discussed the pre-PBAE validation process by the ELBSC on the different components of the test battery, namely Classroom Language Assessment, Speaking, Writing, Reading and Listening. The work of the ELBSC contributed to the development of the Framework for Pilot Benchmark Assessment (English Language), which states the purposes, format and the structure of the PBAE. Chapter 7 describes the Pilot Benchmark Assessment phase of the consultancy study.

References

Coniam, D. (2001). The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: A case study. *SYSTEM, 29*(2), 1–14.

Coniam, D., & Falvey, P. (1996). *Setting language benchmarks for English language teachers in Hong Kong secondary schools*. Hong Kong: Advisory Committee on Teacher Education and Qualifications.

Coniam, D., & Falvey, P. (1998a). *Validating the classroom language assessment component: The Hong Kong English language benchmarking initiative*. Hong Kong: Advisory Committee on Teacher Education and Qualifications.

Coniam, D., & Falvey, P. (1998b). *Validating the reading test: The Hong Kong English language benchmarking initiative*. Hong Kong: Advisory Committee on Teacher Education and Qualifications.

Coniam, D., & Falvey, P. (1998c). *Piloting the multiple-choice cloze test: The Hong Kong English language benchmarking initiative*. Hong Kong: Advisory Committee on Teacher Education and Qualifications.

Coniam, D., & Falvey, P. (1998d). *Validating the speaking test: The Hong Kong English language benchmarking initiative*. Hong Kong: Advisory Committee on Teacher Education and Qualifications.

Deyes, T. (1984). Towards an authentic 'discourse cloze'. *Applied Linguistics, 5*(2), 128–137.

Falvey, P., & Coniam, D. (1998a). *Pre-pilot exercise for the rewriting and speaking components of the English language benchmark project*. Hong Kong: Hong Kong Examinations Authority.

Falvey, P., & Coniam, D. (1998b). *Assessor training and standardisation for classroom language assessment: The Hong Kong English Language Benchmarking Initiative*. Hong Kong: Advisory Committee on Teacher Education and Qualifications.

Falvey, P., & Coniam, D. (1999c). *Assessor training and standardisation for the speaking test: The Hong Kong English Language Benchmarking Initiative*. Hong Kong: Advisory Committee on Teacher Education and Qualifications.

Falvey, P., & Coniam, D. (1999d). *Assessor training and standardisation for the writing test: The Hong Kong English Language Benchmarking Initiative*. Hong Kong: Advisory Committee on Teacher Education and Qualifications.

Falvey, P., & Coniam, D. (2000). Establishing English language writing benchmarks for primary and secondary teachers of English language in HongKong.HongKong Journal of Applied Linguistics, 5(1), 128–159.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. Edward Arnold: London.

Nevo, D. (1999). Paper presented at Language Benchmarks Colloquium. September 29, 1999. Hong Kong: Hong Kong Institute of Education.

David Coniam is Chair Professor and Head of Department of the Department of Curriculum and Instruction in the Faculty of Education and Human Development at The Education University of Hong Kong, where he is a teacher educator, working with teachers in Hong Kong primary and secondary schools. His main publication and research interests are in language assessment, language teaching methodology and computer assisted language learning.

Peter Falvey is a teacher educator, formerly a Head of Department in the Faculty of Education, The University of Hong Kong. His main publication and research interests are in language assessment, first and second language writing methodology, and text linguistics.