Illustrating the Demands of a Performance Test Through an Interactive Multimedia Syllabus

David Coniam The Chinese University of Hong Kong

Abstract

This article describes the rationale for and production of a syllabus (or 'test specification') document implemented via hyperlinked multimedia. The purpose of the multimedia program is to illustrate and exemplify the scales, descriptors and criteria for the Classroom Language Assessment (CLA) component of the English Language Benchmark Test in Hong Kong – a newly introduced language certification test for teachers of English as a Second Language (ESL). The CLA test is a performance test of EFIgure 1nglish, and is assessed on the job; that is, as ESL teachers are teaching a live English lesson.

The article first examines the position of printed syllabus documents generally, i.e., within the field of ESL. It then outlines the problems associated with a print syllabus that purports to specify the demands and requirements of an oral performance test. The article then describes how digitized videotaped samples of actual teacher classroom performance have been integrated with the CLA scales and descriptors through a series of HTML pages that run under a Web browser. The multimedia program also contains comments on each videotaped clip that explain why particular video clips exemplify specified levels. Although the multimedia program has been developed essentially for the purpose of exemplifying the demands of a specific test, the program has broader potential in that the suite can obviously also be used for training purposes - standardizing examiners who will be conducting assessments of candidates taking the CLA test.

Vol. 33, No. 2 2001

Background

The issues discussed in this article center around an initiative by the Government of the Hong Kong Special Administrative Region (HKSAR) to establish minimum language proficiency examinations (also known as 'language benchmarks') for all teachers in Hong Kong primary and secondary schools. The genesis of these benchmark examinations lies in concern expressed, since the early 1990s, by different sectors of the business and education communities in Hong Kong over perceived falling language standards. In 1996, as a consequence of these concerns and with a view to upgrading teachers' language proficiency, the establishment of minimum standards tests of language proficiency for teachers in Hong Kong began to be investigated. The first language benchmark test was administered in early 2001 to teachers of English as a Second Language (ESL). The English language benchmark test consists of a battery of 'formal' tests (i.e., Reading, Writing, Listening, Speaking) and a live performance test—the Classroom Language Assessment test (CLA). Although the CLA is the focus of the current article, it should be noted that the article does not address language assessment issues, but rather how the demands of this new test may be communicated to test takers. For details of the background to the research and development of the English language benchmark initiative, readers are referred to articles by Coniam and Falvey, 1996; Falvey and Coniam, 1997, Coniam and Falvey, 1999)1. While the thrust of the article is not on assessment per se, some background to the CLA test is necessary so that the reader may understand the rationale for the multimedia program described herein.

The Classroom Language Assessment Test In the CLA component of the English language benchmark test, English language teachers are assessed (twice by two different examiners for the sake of fairness and reliability) teaching a live lesson. They are evaluated on four scales:

- 1. Grammatical accuracy
- 2. Pronunciation, stress and intonation
- The language of classroom interaction (the pragmatic and strategic areas of appropriate classroom interaction involving register, appropriate 'tailoring' of language to different types and levels of student, eliciting responses from, responding to and providing feedback to students)
- 4. The language of classroom instruction (involving formal explanation and presentation, the framing and checking functions of giving instructions etc)

The first two scales – Grammatical accuracy and Pronunciation, stress and intonation – are the two 'formal' elements that underpin an ESL teacher's language proficiency in English. English language teachers are quite familiar with assessment scales such as these since they appear in many oral tests of English language. Teachers are used to assessing their students on such scales, and are themselves used to being assessed in the same way when taking tests of spoken English administered by public examination bodies.

Syllabi, which are critical to the delivery of appropriate content, are all too often for secondary school teachers simply the table of contents of the course's text.

The other two scales – The language of interaction and The language of instruction – are the 'functional' indicators of a teacher's proficiency in English in terms of communicating with students and getting things done in the classroom. These scales assess language and not pedagogy: they are, however, aspects of assessment that are new to Hong Kong ESL teachers. Subsequently, there has been considerable debate on how the demands posed by these scales should be interpreted.

Each of the four scales consists of five levels. As can be seen from Table 1, the mid-point – level '3' – is 'at the benchmark'.

Table 1: Hong Kong benchmark levels

Level	Gloss
5	Well above the benchmark
4	Above the benchmark
3	At the benchmark
2	Below the benchmark
1	Well below the benchmark

This article is primarily concerned with making the demands of a syllabus—or test specification—document² more transparent and more accessible to test takers. A brief discussion of print syllabi—for both testing and teaching purposes—will therefore follow to provide the reader with a broader perspective on the accessibility and usability of syllabus documents. This will then be followed by details of the implementation of the Hong Kong CLA multimedia syllabus.

Accessibility of Syllabus

documents

In Hong Kong, for example, many teachers have never looked at either the school teaching syllabus or the public examination syllabus. Syllabi, which are critical to the delivery of appropriate content, are all too often for secondary school teachers simply the table of contents of the course's text. A number of researchers have examined the extent to which the textbook dominates classroom practice (see e.g., Ball and Feiman-Nemser, 1988; Cheng, 1997; Woodward and Elliot, 1990).

In part, this may be because syllabus documents are often produced by bodies of 'experts'—such as the Education Department in Hong Kong. Although these bodies may attempt to make the syllabus document as user-friendly as possible, the descriptions of syllabus principles and even methodological practices will inevitably be couched more in the language of the expert, the syllabus designer, rather than in the language of the teacher.

One consequence of the high technical complexity of the syllabi is that some teachers rarely consult them. In Hong Kong, for example, many teachers have never looked at either the school teaching syllabus or the public examination syllabus. The teaching syllabus is produced by the Hong Kong Education Department and is intended to give a purposeful sense of direction to the teaching of ESL. The public examination syllabus is produced by the Hong Kong Examinations Authority: it is intended to interface with the teaching syllabus to give the examinations content validity, and specify what knowledge and skills will be assessed. To many teachers, then, the teaching syllabus is simply the textbook (and the examination syllabus is simply past examination papers). Reasons for this may be that teachers feel the textbook is all they need or all they have time for in a full working life within the school system. It may be because they view themselves as practitioners and find that the manner in which the syllabus is written, or the material described in the syllabus, is too removed from, or does not match with, their practice.

In the context of a religious education syllabus, Malone (2000) describes how aspects of that syllabus were poorly implemented because certain teachers did not understand the approach as it was laid out in the curriculum documents. She cites the "theoretical dissonance" (Malone 2000, 3) between the curriculum 'model' and how this translated into classroom practice. She notes that the majority of teachers had not integrated the model into their teaching, and essentially approached the curriculum support units "as a source of activities to keep their students occupied". (Malone 2000, 3)

Accessibilityof The CLA Test Syllabus

For teachers whose test preparation for their students involves not referring to the syllabus, but essentially relying on past examination papers, the new CLA test looms rather forbiddingly.

One of the objectives of the current multimedia implementation of the CLA syllabus is therefore to make higher-level concepts transparent for both test takers and examiners, and to illustrate how these concepts translate into demands made upon test takers in the actual test situation.

The accessibility problems discussed above concerning printed syllabus documents are compounded in the syllabus for the CLA component of the English language benchmark test. The HKSAR Government has striven to produce a detailed syllabus: the published document contains the scales and descriptors against which test takers will be assessed and contains as many exemplars of the different test types as was possible to put together in the limited time frame prior to publication of the benchmark test specification document. In a print document alone, however, it is not easy to convey to test takers the exact demands of a performance test such as the CLA component.

With regard to the formal skills component tests of the English language benchmark test – i.e., the Reading, Listening and Writing Tests – teachers can be expected to be reasonably familiar with the general demands posed by these tests, since they prepare and administer tests of a broadly similar nature to their own students. However, as mentioned above, it is difficult for teachers to fully appreciate the demands of a performance test such as the CLA from textual description and explanation alone, without actual samples of performances exemplifying the assessment scales at different levels of ability (see Griffin and Nix, 1991). For this reason, many teachers are alarmed about what is expected of them in the CLA test. For teachers whose test preparation for their students involves not referring to the syllabus, but essentially relying on past examination papers, the new CLA test looms rather forbiddingly.

Consider for example, the benchmark scale for Pronunciation, stress and intonation, and the descriptor of performance at the benchmark level:

Pronunciation of sounds is generally acceptable although there are some errors in the pronunciation of sounds and/or word-stress and a number of L1 characteristics are evident but are not obtrusive. Sentence stress and intonation patterns may sometimes be inappropriate but communication is seldom impeded. (Gov't of the HKSAR 2000, 9)

It is instructive to consider how a teacher might interpret some of the crucial points in the above descriptor:

- "... generally acceptable" Does this imply that errors are permitted? If so, how many?
- " ...a number of L 1 characteristics are evident but are not obtrusive" How is the term "obtrusive" to be interpreted?
- "... but communication is seldom impeded"

The phrase above is glossed in the syllabus document as follows:

Because of long exposure to the teacher and familiarity with L1 interference in L2 speech, the students may not always experience undue strain or difficulty in understanding what the teacher says. However, as a classroom model of Hong Kong educated spoken language, the teacher's Pronunciation, stress and intonation are unacceptable because they are likely to perpetuate unacceptable levels of Pronunciation, stress and intonation in the students. (Gov't. of the HKSAR 2000, 10)

To an examiner who has undergone thorough training and who has observed teachers teaching on video and in the live classroom, the above descriptor can be interpreted sufficiently well to make a reasoned assessment. To a teacher, however, what constitutes the benchmark level is much less apparent, and many teachers are apprehensive about whether their own Pronunciation, stress and intonation will be assessed as an 'unacceptable' model or not.

The Use and Advantages of Multimedia

'Interactivity'

An important principle advocated by educational technologists regarding the use of technology – both in the classroom and for self-access purposes – is that users should not simply watch, but need to construct knowledge (Jonassen, 1995; 1996). The constructivist theory that follows from this is that students need to actively 'participate' in technology, and that by so doing will adjust their world view, their knowledge base. This call is echoed in some of the principles accompanying discovery learning (see Jacobs (1992), Goforth (1994) for a discussion of the use of discovery learning aided by technology).

In the field of language and ESL in particular, multimedia is beginning to be used extensively. It tends to be used mainly for exemplifying language concepts rather than in the form of a syllabus document: see e.g., English Computerized Learning (1996); Mueller and Samson (1998), where video is incorporated with explanations of the articulations of English phonemes.

What might be termed 'enriched input' provided by interactive lessons on CD-ROM is also beginning to appear as an adjunct to teacher education and teacher development programmes (see e.g., Jiva Institute, 2000).

In the field of ESL, Johnson's (1998) multimedia package "Teachers Understanding Teaching" illustrates admirably how to get teachers to critically reflect on and evaluate their own beliefs and classroom practices. The CD-ROM consists of a series of carefully-selected clips of both classroom teaching and interviews with ESL teachers in which they reflect and discuss. As well as activities for participants to consider as they watch, the package also consists of a series of extension activities for teachers to conduct upon themselves and their peers both inside and outside their own classrooms to further the critical reflection that Johnson is attempting to promote. Constructivist outcomes can clearly be discerned in the package.

With regard to the multimedia test syllabus described in the current article, then, users (i.e., teacher examinees) will ideally construct an understanding of what the CLA test aims to assess, and will arrive at an understanding of how to best meet the assessment requirements. While the current program aims to encourage teachers to reflect on their own proficiency, and to relate this more concretely to the syllabus, it must be acknowledged that the program does not have as much interactive potential as Johnson's (1998) package. Nonetheless, the set-up of the current program is an attempt to promote more than simply passive watching. As regards 'interactivity', however, the fact that the CLA syllabus consists of a fairly limited set of choices means that the term 'interpassive' is, on balance, a better categorization of the multimedia CLA syllabus.

As with any test of performance, assessment personnel need to be trained and standardized for the Classroom Language Assessment component. In Hong Kong, training has, in the past, centered around the use of videotaped lessons, the use of which raises a number of issues. Firstly, an analogue medium such as video is cumbersome to store and use. Secondly, videos are a much less user-friendly medium than CD-ROMs. It is difficult to move back and forth in videos in order to locate / replay crucial segments so that comparisons between different aspects of teacher performance may be made. With digitized video clips, it is possible to move from one clip to another at the click of a mouse. Consequently, with CD-ROMs, judgments can be honed and refined more precisely. Thirdly, videos deteriorate over time; after a few years they often have to be discarded.

As regards 'interactivity', however, the fact that the CLA syllabus consists of a fairly limited set of choices means that the term 'interpassive' is, on balance, a better categorization of the multimedia CLA syllabus.

Advantages of a Multimedia Implementation

Program Production A description of the process involved in the production of the CD-ROM-based syllabus from samples of classroom teaching will now follow, as well as a description of the workings of the multimedia program itself.

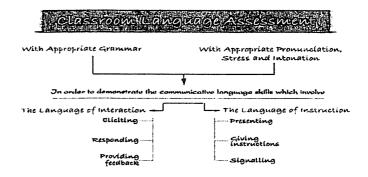
A number of live English language lessons (approximately 45 forty-minute lessons) were video-recorded (with the teachers' permission), and the videos digitized into MPEG-1 format. The clips were then assessed by trained staff from the HKSAR Education Department's English Language Benchmark Assessment Unit, who had undergone substantial training and standardization in assessing teachers for the benchmark test (Falvey and Coniam, 1999).

The videos were transcribed and analyzed qualitatively for segments that matched the criteria in the CLA syllabus. The assessors' comments were also recorded and transcribed.

A program template was developed, consisting of a series of hyperlinked HTML pages, using Microsoft Front Page (Microsoft, 2000). An implementation in HTML obviously permits links to be easily made between elements. In the current project, these are principally between the scales and the descriptions of the benchmark levels to video clips, transcripts and comments.

After passing through some introductory screens that outline such matters as confidentiality, the opening screen is an adapted graphic from the Syllabus Specifications document (Government of the HKSAR, April 2000).

Figure 1: Opening screen - the four scales



The layout of the diagram in Figure 1 matches the

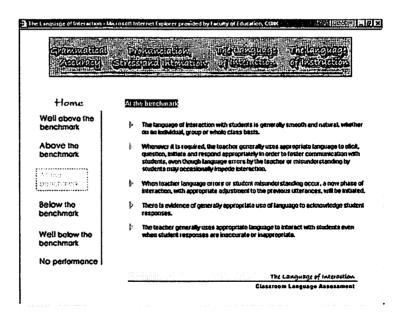
Navigating within the program

conceptualization - described earlier - of the scales in terms of two 'formal' and two 'functional' scales. The links to the two scales at the top of the diagram are to the scales and descriptors for Grammatical accuracy (shown as "With Appropriate Grammar") and Pronunciation, stress and intonation (shown as "With Appropriate Pronunciation, Stress And Intonation"). Beneath these two formal scales are the two functional scales (The language of instruction and The language of interaction).

As mentioned, the program has been constructed as a series of hyperlinked HTML pages. If a user clicks on the bottom left scale The Language of interaction, she is taken to the full The Language of interaction scale, which has been adapted from the Syllabus Specifications document.

The left-hand frame contains the different benchmark levels, while the central frame of the web page contains the criteria for the particular level.

Figure 2: Criteria for The Language of Interaction scale, At the benchmark



In Figure 2, it will be seen that 'At the benchmark' has been selected, and is highlighted, or boxed. The criteria to the right

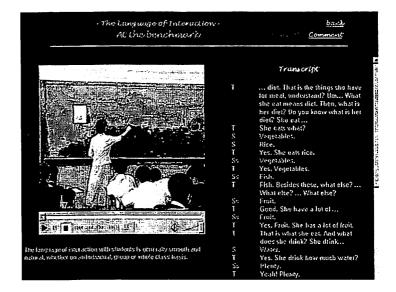
The comments are an important feature of the program in that they attempt to put the scales and descriptors into perspective by identifying the linguistic features that determine the level of a particular clip.

hand side of the web page are for this level only.

In this screen, the left-hand frame contains the different benchmark levels. Clicking on a given level then brings up the relevant set of criteria for the particular scale and level.

If a user now clicks on one of the criteria in the central hand frame, she is taken to a page containing a related lesson clip. Figure 3 illustrates.

Figure 3: The Language of Interaction, At the benchmark

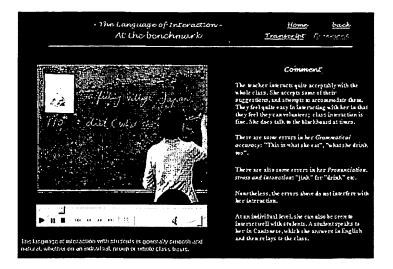


When the page appears, its video clip plays automatically. The page also contains, at the bottom left, the criterion that a particular video clip is demonstrating. A transcript of the lesson segment appears to the right-hand side of the page, and the 'Transcript' button (top right) is highlighted. The purpose of the transcript is to enable the lesson segment – which the clip is demonstrating - to be more easily followed; at times, this can be difficult because the teacher may be talking quietly or talking to a student. The transcript consequently enables the viewer more easily to work out what is taking place in the lesson.

If the user clicks on the 'Comment' button (top right – beside the 'Transcript' button), a comment – drawn from a qualitative

analysis of the particular lesson by experienced assessors – appears in place of the transcript.

Figure 4: Comment and video clip for The Language of Interaction, At the benchmark



Users have stated that the nature and set-up of the program not only allow them to understand what is expected of a benchmark level performance, but have also enabled them ... to go beyond the program, and to reflect upon their own experience and proficiency...

The comments are an important feature of the program in that they attempt to put the scales and descriptors into perspective by identifying the linguistic features that determine the level of a particular clip. As with Johnson's (1998) CD-ROM for teacher development, it is the inclusion of material such as evaluative comments that will encourage users to reflect on what they are watching, and to go beyond the clip as they hopefully relate the content to themselves and to their own teaching situation.

The clip in Figure 4 above, for example, shows a teacher who is 'at the benchmark' for The language of interaction. In the clip, the teacher makes a number of errors of grammar and pronunciation. Part of the comment is therefore taken up with explaining that such errors should not impinge on how the teacher is assessed on The language of interaction, where she meets the benchmark. While such comments obviously relate to the video clip being observed, the comments also serve as a springboard in assisting teachers to reflect upon their own language proficiency in the different areas of classroom language.

The program has been pilot-tested on four groups of English

Discussion and Conclusion

language teachers (approximately 100) who will soon be facing the CLA test. Their reaction has been positive, with many commenting that it offers a perspective on the CLA component of the benchmark test that far surpasses the print document. Users have stated that the nature and set-up of the program not only allow them to understand what is expected of a benchmark level performance, but have also enabled them (as more active participants) to go beyond the program, and to reflect upon their own experience and proficiency — which is not possible in a static print document.

This article has described a teacher support project aimed at making a print-based assessment syllabus more accessible, user-friendly and ultimately understandable. The manner in which the project has been undertaken has not merely involved moving a print document from a text medium to the computer. The use of hypertext makes any long document – even one consisting simply of text – more accessible by virtue of the fact that overt content links can be specified. In the presentation of the requirements of an oral performance test such as the Hong Kong CLA benchmark test, more than textual descriptions are necessary.

The purpose of the project has therefore been to create an instructional support document in a multimedia format so that it is accessible and understandable as a syllabus document as well as utilizing the possibilities which multimedia offers. This means that while video samples have been used to exemplify criteria, other features such as transcripts of video clips for better understanding of the language therein have also been included, as well as qualitative comments to support the assessment made of the contents of a particular clip. The program can be used independently (or with small groups of teachers) as a tool not only for appreciating the demands of the CLA test but also for gauging one's own level against the exemplars.

The production of such a program has not been a small task. A great deal of planning and effort has been necessary at all stages in the development of the program. A considerable amount of time was spent working at each stage, from the initial phase of conceptualizing how the 'information' in the syllabus document might be best presented to the final phase of how the technical 'detail' (i.e., the test criteria, scales and descriptors) might be molded into a usable instrument. One time-consuming aspect in the development of the project has been the videotaping of an appropriate sample of teachers, and the

accompanying time spent in transcribing and evaluating the lessons and subsequently in identifying clips that exemplify the relevant scales and descriptors.

While the focus of this article has involved the implementation of a multimedia syllabus in order to render a print-based syllabus document for a performance test more accessible and less daunting, the product itself – as well as the conceptual backdrop to the project – has potential in several ways. As a syllabus document in itself, the use of multimedia is a powerful way to make productive tests more 'accessible' in that it makes it easier for potential test takers to understand the demands of the test and therefore how to prepare for it. As regards the Hong Kong English language benchmark test, apart from the Classroom Language Assessment, a Speaking Test also forms part of the benchmark test battery (Government of the HKSAR, 2000). A multimedia CD-ROM version along the lines of the current project for CLA would be welcomed by teachers, as the demands and requirements of the different tasks in that test would be rendered more transparent.

As institutions for teacher education in Hong Kong continue to conduct benchmark assessment, a bank of exemplar video clips will become a useful tool in training and standardizing staff at the different institutions. With what is, in effect, an easily accessible database of teacher performance at various levels, it will be much more convenient for teaching staff in the teacher education institutions to train and standardize their staff along similar lines. It will be possible to show exactly what the benchmark level is, rather than allowing teacher education institutes to come to their own conclusions as to what the scales and descriptors demand and how to interpret what constitutes a benchmark level 'pass' or 'fail'.

Finally, while this article has described a multimedia implementation of a CD-ROM targeted at clarifying the principles involved in the assessment of language benchmarks for teachers, the multimedia program format could easily be adapted to exemplify strategies and techniques for teaching ESL rather than for assessing it. The assessment scales might, for example, be replaced by headings such as "Teaching reading", "Teaching writing" etc. or "Classroom management techniques". How different aspects of these skills might be successfully taught in the classroom could then be linked in to video clips and comments that illustrate and outline good practice. As either a classroom-based aid or as a discovery-learning tool for English language teachers in training, multimedia can be a

powerful aid in raising awareness and conceptualization in both teaching and learning. •

Notes

- ¹ There are approximately 42,000 primary and secondary school teachers, of whom approximately 12,500 are English language teachers. Benchmarks are being investigated on two fronts. The first concerns language teachers, that is, teachers of English, Chinese and Putonghua. The second concerns teachers who teach content subjects (history, geography, biology, mathematics etc.) through the mediums of either English or Chinese (see Falvey and Coniam, 1997).
- ² The term 'syllabus document' refers to the test specification document produced in Hong Kong by the body responsible for public examinations, the Hong Kong Examinations Authority. The document details the test objectives, that is, what the different test components are intended to assess, as well as providing specimen test papers and marking schemes.

Works Cited

- Ball, D. and S. Feiman-Nemser, 1988. Using textbooks and teachers guides: A dilemma for beginning teachers and teacher educators. Curriculum Enquiry, 18, 401-423.
- Cheng, L.Y., 1997. The washback effect of public examination change on classroom teaching. Unpublished PhD thesis, Hong Kong: The University of Hong Kong.
- Coniam, D. & P. Falvey, 1996. Setting language benchmarks for English language teachers in Hong Kong secondary schools. Hong Kong: Advisory Committee on Teacher Education and Qualifications.
- Coniam, D. & P. Falvey, 1999. The English language benchmark initiative: A validation study of the Classroom Language Assessment component. Asian Pacific Journal of Language in Education, 2 (2), 1–35.
- Dufresne, R.J., W.J. Gerace, W.J. Leonard, J.P. Mestre, , & L. Wenk, 1996. Classtalk: A classroom communication system for active learning. Journal of Computing in Higher Education, 7, 3-47.
- English Computerized Learning Inc., 1996. Pronunciation power. Edmonton: Alta (http://www.EnglishLearning.com).

- Falvey, P. & D. Coniam, 1997. Introducing English language benchmarks for Hong Kong teachers: a preliminary overview. Curriculum Forum, 6 (2), 16-35.
- Goforth, D., 1994. Learner Control = Decision Making + Information: a Model and Meta-analysis. Journal of Educational Computing Research, 11 (1), 1-26.
- Government of the Hong Kong Special Administrative Region (HKSAR), 2000. Language Benchmark Assessment for Teachers English Language: Syllabus Specifications, Explanatory Notes, Specimen Questions with Suggested Answers, Scales and Descriptors. Hong Kong: Government Printer.
- Griffin, P. & P. Nix, 1991. Educational assessment and reporting: a new approach. Marrickville, NSW: Harcourt Brace Jovanovich.
- Jacobs, G., 1992. Hypermedia and discovery-based learning: A historical perspective. British Journal of Educational Technology 23 (2), 113-121.
- Jiva Institute, 2000. Teacher training CD-ROM. http://www.jiva.org/Education/JLS/jlsteacher.htm (26 June 2000).
- Johnson, K. E., 1998. Teachers Understanding Teaching. Boston, MA: Heinle & Heinle.
- Jonassen, D.H., 1995. Computers as cognitive tools: Learning with technology and not from technology. Journal of Computing in Higher Education, 6, 40-73.
- Jonassen, D. H., 1996. Computers in the classroom: Mindtools for critical thinking. Englewood Cliffs, NJ: Prentice-Hall.
- Microsoft Corporation, 2000. Front Page 2000. http://www.microsoft.com/.
- Mueller, A & E. Samson, 1998. Pronunciation and speech tutor. Hong Kong: University of Hong Kong.
- Woodward, A. & D. Elliott, 1990. Textbooks: Consensus and controversy. In Woodward, A. and D. Elliott, eds. Textbooks and Schooling in the United States. 89th Yearbook of the National Society of the Study of Education. Chicago: University of Chicago Press.

David Coniam is a Professor in the Faculty of Education at The Chinese University of Hong Kong, where he is a teacher educator, working with ESL teachers in Hong Kong secondary schools. His main publication and research interests are in language testing, computational linguistics and language teaching methodology.