



**CAMBRIDGE ENGLISH**  
Language Assessment  
Part of the University of Cambridge

**100** CAMBRIDGE  
**ENGLISH**  
CENTENARY 1913–2013

# Research Notes

Issue 51

February 2013



ISSN 1756-509X



**CAMBRIDGE ENGLISH**  
**Language Assessment**  
Part of the University of Cambridge

# Research Notes

Issue 51 / February 2013

A quarterly publication reporting on research, test development and validation

## *Senior Editor and Editor*

Dr Hanan Khalifa, *Head of Research and International Development*, Research and Validation Group, Cambridge English Language Assessment

Coreen Docherty, *Senior Research and Validation Manager*, Research and Validation Group, Cambridge English Language Assessment

## *Editorial Board*

Dr Ardeshir Geranpayeh, *Head of Psychometrics and Data Services*, Research and Validation Group, Cambridge English Language Assessment

Nigel Pike, *Head of Assessment*, Assessment and Operations Group, Cambridge English Language Assessment

Dr Nick Saville, *Director*, Research and Validation Group, Cambridge English Language Assessment

Ron Zeronis, *Assessment Group Manager*, Assessment and Operations Group, Cambridge English Language Assessment

## *Production Team*

Rachel Rudge, *Production Controller*, Cambridge English Language Assessment

John Savage, *Publications Assistant*, Cambridge English Language Assessment

Printed in the United Kingdom by Océ (UK) Ltd.

# Research Notes

## Contents

<b>Measured constructs: A history of Cambridge English language examinations 1913–2012</b> Cyril J Weir	<b>2</b>
<b>Components of an elaborated approach to test validation</b> Gad S Lim	<b>11</b>
<b>A look at test taker characteristics for a C2 exam</b> Ardeshir Geranpayeh	<b>15</b>
<b>Consulting stakeholders as part of the Cambridge English: Proficiency exam revision</b> Coreen Docherty and Debbie Howden	<b>18</b>
<b>Development and construct of revised Cambridge English: Proficiency</b> Ron Zeronis and Mark Elliott	<b>22</b>
<b>A look into the future</b> Michael Milanovic	<b>31</b>
<b>2012 Research Notes reader survey</b> Fiona Barker and Coreen Docherty	<b>34</b>
<b>Caroline Clapham IELTS Masters Award 2012</b>	<b>37</b>
<b>Studies in Language Testing</b>	<b>38</b>
<b>ALTE report</b>	<b>39</b>

## Editorial notes

Welcome to issue 51 of *Research Notes*, our quarterly publication reporting on matters relating to research, test development and validation within Cambridge English Language Assessment, the new name for Cambridge ESOL. This issue celebrates 100 years of Cambridge English examinations, as well as the launch of the new *Cambridge English: Proficiency* exam, by looking at the past, present and future of the exam board and this groundbreaking exam.

In the first article, Professor Cyril J Weir provides a historical account of the development of Cambridge English language examinations. He describes the evolution of the exam board and how major changes have been made in response to key events in the field of language learning and assessment at various points in history. This is followed by Gad S Lim's discussion of how Cambridge English Language Assessment has refined its approach to test validation (VRIPQ) over the last decade.

The next three articles focus on the test development process involved in updating the new *Cambridge English: Proficiency* exam, which was undertaken between 2009 and 2012. Ardeshir Geranpayeh provides a report on the characteristics of *Cambridge English: Proficiency* candidates between 1991 and 2012, the results of which contributed to the revision of this exam. This is followed by Coreen Docherty and Debbie Howden's report on an investigation into stakeholder perceptions of the new exam during two different stages of the revision process. Then, Ron Zeronis and Mark Elliott provide a detailed description of the new *Cambridge English: Proficiency* exam and the rationale behind the changes made to it. These articles highlight the importance of having an explicit and rational approach to test development and how Cambridge English Language Assessment has applied the principles of 'impact by design'.

Next, Michael Milanovic, the Chief Executive of Cambridge English Language Assessment, looks to the future and describes the role Cambridge English Language Assessment may play in the next 100 years.

Finally, although *Research Notes* has not reached its centenary, we are pleased to report the results of a reader survey conducted after reaching its 50th issue. We are also pleased to announce the 2012 Caroline Clapham IELTS Masters Award winner, provide details of the most recent volumes of the *Studies in Language Testing* series to be published and Martin Nuttall, from the ALTE Secretariat, reports on recent and upcoming ALTE activities.

# Measured constructs: A history of Cambridge English language examinations 1913–2012

CYRIL J WEIR POWDRILL CHAIR IN ENGLISH LANGUAGE ACQUISITION, CRELLA, UNIVERSITY OF BEDFORDSHIRE, UK

## A 100-year journey

In recent years assessment issues have assumed increased importance in the economic, educational and socio-political affairs of society. Spolsky (2008:297) argues that 'testing has become big business', and Shohamy (2008:xiv) points to 'the societal role that language tests perform, the power that they hold, and their central functions in education, politics and society'. A significant role for testing language proficiency can be seen *inter alia* in migration and citizenship policy and practice, the professional registration of those involved in the provision of health care, appointment and promotion in business, industry and commerce, the certification of air traffic and maritime personnel, and entry to tertiary level education. Such uses testify to the critical function that language assessment now fulfils in contemporary society.

As the power of tests and the potential for their misuse/abuse grows, assessment literacy seems more important than ever. Taylor (2009:29) has suggested that narrative accounts which chronicle testing developments over time may have an important role to play in fostering this:

They contextualize the practice of language testing as a socially constructed and interpreted phenomenon, rather than treating it primarily as a pseudoscientific endeavour that is removed and isolated from human individuals and social values. It may well be that popular adaptations of this narrative, storytelling approach will prove a more effective means of developing assessment literacy among the wider stakeholder community in the future.

Despite the contemporary importance of language testing in the United Kingdom, we have always lacked a satisfactory account of its historical development in this country. In *A History of English Language Teaching*, Howatt and Widdowson (2004:332) acknowledged that they had not given English language testing 'the prominence it deserves'. Spolsky's (1995) authoritative work on language testing, *Measured Words*, for the most part covered the development of English language testing in the United States with only partial reference to events on this side of the Atlantic. The recent *Studies in Language Testing* (SiLT) volume *Measured Constructs* (Weir, Vidaković and Galaczi forthcoming 2013) sought to make good this deficit by adopting a mainly British perspective, focusing on Cambridge English language examinations over the course of the last century. This took the reader from a small cottage industry in 1913, with the Certificate of Proficiency in English (CPE) (now known as *Cambridge English: Proficiency*) administered to three students in one London centre, to the big business of a leading international examining board in 2012. Cambridge now offers multiple English language examinations at different levels, across different domains, to nearly 4 million candidates per annum in 2,700 authorised centres across 130 countries (based on 2011 figures).

In the final chapter of Weir and Milanovic ((Eds) 2003), the SiLT volume on the CPE 2002 revision, Weir emphasised the need for Cambridge to continue research into the complex cognitive processes and attendant performance conditions involved in completing the tasks in its English language examinations. This was to be one of the core aims of the set of four 'constructs' volumes in the SiLT series (*Examining Writing* (Shaw and Weir 2007), *Examining Reading* (Khalifa and Weir 2009), *Examining Speaking* (Taylor (Ed.) 2011) and *Examining Listening* (Geranpayeh and Taylor (Eds) forthcoming 2013), which provided an informed, comprehensive, synchronic analysis of contemporary Cambridge English examinations. *Measured Constructs* (Weir et al forthcoming 2013) has continued with that endeavour but differs in that it takes a diachronic approach. It provides further insight into the constructs being measured by investigating how the testing of each macro language skill has evolved over the last 100 years.

*Measured Constructs* provides a synopsis of the powerful influences that language teaching had on general English language examinations for non-native speakers developed at Cambridge from 1913–2012. Most attention is paid to the CPE between 1913 and 2012 and the Lower Certificate in English (LCE) between 1939 and 2012, rebranded after 1975 as the First Certificate in English (FCE) and now known as *Cambridge English: First*, as these offered an accessible and manageable historical perspective on Cambridge English examinations over an extended period of time. Such a focus allowed the authors to trace *continuity* and *innovation* in the measurement of language constructs in one examination board over 100 years of its history.

The authors began by identifying the pedagogic legacies from the past that affected the first CPE examination in 1913 (Stage 1) which can be seen as a hybrid creation influenced by:

1. The **Grammar-Translation Method** (Meidinger 1783, Fick 1793), reflected in the inclusion of translation tasks and questions on grammar. The Grammar-Translation Method was originally an attempt to adapt the scholastic study of foreign languages for a reading knowledge of their culture and history 'to the circumstances and requirements of school students' (Howatt 1984:131). Throughout the 19th century, proponents of the Grammar-Translation Method tried to carve out a role for it in teaching modern languages in the schools by modelling their classroom procedures on the teaching of Latin and Greek with translation of (literary) texts being seen as the main activity in language learning.
2. The **Reform Movement** (Viëtor 1882, Passy 1899, Jespersen 1904), reflected in the inclusion of a phonetics paper, a speaking paper and an essay paper. The Grammar-Translation Method was not popular with some teachers, however, and in the 1880s a number of language teachers and academics in Europe instigated the Reform Movement which, with the assistance of modern ideas from phonetics,

allowed for a new pedagogical approach rooted in the spoken language.

3. The **Examinations in Modern Languages Approach**, which were national, modern language assessments in use at the start of World War One and contained activities derived from both the Reform Movement and the Grammar-Translation Method, as in the French examination in the School Leaving Certificate:

They will be required to answer certain questions on French grammar . . . Moreover they will be required to do two pieces of translation, into and out of French respectively, and a piece of free composition in French, and finally to submit to a short oral test of their ability to read aloud, write dictation and converse in French (Palmer and Redman 1932:80).

**Table 1: 1913 CPE examination**

<b>Written:</b>	(a) Translation from English into French or German	2 hours
	(b) Translation from French or German into English, and questions on English grammar	2½ hours
	(c) English Essay	2 hours
	(d) English Literature (The paper on English Language and Literature [Group A, Subject 1] in the Higher Local Examination)	3 hours
	(e) English Phonetics	1½ hours
<b>Oral:</b>	(a) Dictation	½ hour
	(b) Reading and Conversation	½ hour

As Table 1 shows, the written part of the examination included an English Literature paper and an English Essay, but it also contained a compulsory English Phonetics paper, a translation task from and into French or German, and an English grammar section. There was also an oral component with dictation, reading aloud and conversation. There was a focus on form in the grammar and phonetics sections in the first CPE but attention was clearly paid to active language use as well. In all, a demanding 12 hours of testing against the fewer than 4 hours required at *Cambridge English: Proficiency* today (see the Appendix for a copy of the 1913 examination).

Marks were distributed as in Table 2 below (from a note discovered in the Cambridge Assessment archives among the personal papers of Flather, the Secretary of the Syndicate (1910–21)):

**Table 2: Weighting given to each paper**

Certificates of Proficiency Paper Weightings	
At the Examiners' Meeting held in February 1913 it was proposed that the full marks for each certificate should be 600 to be distributed as follows:	
Phonetics Paper	75
Oral Examination	125 (namely, Dictation 50, Reading and Conversation 75)
Other Papers	100 each
The minimum for passing to be 30% in the Essays, 40% each in Translation, Composition, Phonetics, and Oral; and further that candidates be required to get 200 out of 400 in the whole of the written work except Phonetics taken together, and 100 out of 200 in Phonetics and Oral taken together.	

The CPE was instituted by the Local Examinations Syndicate in 1913 alongside Certificates in Proficiency for teachers in other languages, i.e. French and German. Its purpose is clearly stated in the 1913 *Regulations for the Examinations for*

*Certificates of Proficiency in Modern Languages and Religious Knowledge*:

The Certificate of Proficiency in English is designed for Foreign Students who desire a satisfactory proof of their knowledge of the language with a view to teaching it in foreign schools. The Certificate is not, however, limited to Foreign Students (University of Cambridge Local Examinations Syndicate 1913:5).

This narrowly defined purpose would disappear from the regulations by 1933 and we find that by 1947 CPE was 'open to all candidates whose mother tongue is not English and it is designed not only for prospective teachers but also for other students with a wide range of interest within the field of English studies' (University of Cambridge Local Examinations Syndicate 1947).

Stage 2 of the historical survey examined the influence of the **Oral - Structural - Situational Approaches** to English language teaching that emerged in the United Kingdom between 1921 and 1970. The 1920s saw the birth of 'structural linguistics' which was to have a powerful and long-lasting influence on the theory and practice of language teaching for the next 50 years. In the USA, the structuralist approaches to linguistics of Bloomfield (1926, 1933) and Fries (1945) often resulted in a primary focus on linguistic form in the language classroom (and on language constituents in assessment). The British structural approach, as epitomised in the work of Palmer (1921a, 1921b), differed in that although it was similarly characterised by attention to graded grammatical structures and systematic word lists, these were usually combined with the direct method with its emphasis on the spoken language. This would be complemented later on in this period by the situational approach in the United Kingdom, which sought to locate the teaching of structural items in simple interesting and relevant situations which made their meaning clear.

Finally, the authors traced the effects of the communicative movement (1971–2012) on Cambridge English examinations. In the 1970s and 1980s we can determine a gradual shift in the United Kingdom away from structural approaches to language teaching towards approaches which involved using language as a means of communication. This takes us into Stage 3 of the historical survey: **Communicative Approaches to Language Teaching and Testing**.

In the classroom there was a growing interest in the functional and communicative potential of language, communicative ability rather than knowledge of structures *per se* (see Richards and Rogers 2001:153–177). The developing communicative approach (a) signalled the importance of meaningful activity (i.e. a reaction against mindless drilling), (b) gave birth to the notional-functional syllabus, (c) built on developments in the growing field of sociolinguistics and (d) promoted an interest in authentic materials. Additionally, it was to provide a conceptual framework for a more comprehensive, richer and transparent specification of content for learning and assessment.

To help understand the relationships between the teaching and the testing of the English language in the United Kingdom from 1913–2012, the authors constructed an outline of key events in the period, building on an original suggestion by Tony Howatt (personal communication), which forms the basis for the published history. This overview is reproduced in Table 3.

It offers a useful set of pegs on which to hang elements of ‘a big picture’ of the historical pattern of language teaching and assessment, as well as of specific events in language testing at Cambridge over the period 1913–2012. Phrases like ‘onwards’ are employed (a) to stress the fact that change is not immediate but takes some time to establish itself in classroom pedagogy and assessment, and (b) to act as a reminder of the fact that successful approaches do not mean the end of

existing ones. They exist side by side for as long as they are felt to be useful. Three broad stages were considered:

- **Stage 1 (1780–1913) The Beginnings of Theory**
- **Stage 2 (1921–c1970) Oral-Structural-Situational Approaches to Language Teaching and Testing**
- **Stage 3 (c1971–) Communicative Approaches to Language Teaching and Testing**

**Table 3: Key events affecting the historical development of English language teaching in the United Kingdom and English language testing at Cambridge**

<b>Stage 1 (1780–1913)</b>	<b>The Beginnings of Theory</b>
1780s onwards	The Grammar-Translation Method (Meidinger 1783, Fick 1793)
<b>1858</b>	<b>The University of Cambridge Local Examinations Syndicate (UCLES) founded</b>
1870s onwards	The Direct Method (Berlitz schools 1878–)
1882 onwards	The Reform Movement: (Viëtor – ‘ <i>Quousque tandem?</i> ’ (1882); Sweet (1899) <i>The Practical Study of Languages. A Guide for Teachers and Learners</i> ; Passy’s essay on the direct method (1899) and Jespersen (1904)
1886	Foundation of the International Phonetics Association
1888	Edgeworth’s papers on reliability (1888)
1892	Foundation of the Modern Language Association of Great Britain
<b>1913</b>	<b>The Certificate of Proficiency in English (CPE)</b>
<b>Stage 2 (1921–c1970)</b>	<b>Oral-Structural-Situational Approaches to Language Teaching and Testing</b>
1920s onwards	The Oral Method (Palmer 1921a) aligned with systematic, graded structural progression <i>The Oxford English Course</i> Parts I–IV (Faucett 1933–34) <i>Essential English for Foreign Students</i> (Eckersley 1938–42) The Structural approach (Bloomfield 1926, 1933 and Fries 1945)
<b>1925</b>	<b>J O Roach joins UCLES as Assistant Secretary (until 1945)</b>
<b>1932</b>	<b>CPE: Phonetics paper and grammar knowledge questions disappear</b>
1936	<i>Interim Report on Vocabulary Selection for the Teaching of English as a Foreign Language</i> (Faucett, Palmer, Thorndike and West 1936)
<b>1939</b>	<b>The Lower Certificate in English (LCE)</b>
<b>1941</b>	<b>The UCLES-British Council Joint Committee</b>
1950s onwards	The Situational Approach (Hornby 1950) Pattern Practice (Fries and Lado 1962) <i>The Oxford Progressive English Course</i> (Hornby 1954–56)
<b>1956</b>	<b>CPE: Use of English paper included as an option</b>
<b>1957</b>	<b>UCLES Executive Committee for the Syndicate’s examinations in English for foreign students</b>
1960	The audio-lingual approach (Brooks 1960)
<b>1960</b>	<b>Lado’s visit to UCLES 1960 in a personal capacity to discuss testing matters</b>
<b>1966</b>	<b>Wyatt (Secretary of the Syndicate 1961–72) visits ETS Princeton</b> <b>CPE revision: availability of a language-only pathway</b> <b>CPE: Use of English Paper, 3-option multiple-choice items introduced</b>
<b>1970</b>	<b>LCE: Structure and Usage paper</b>
<b>Stage 3 (c1971–)</b>	<b>Communicative Approaches to Language Teaching and Testing</b>
1971 onwards	Rüschlikon Symposium 1971; Council of Europe initiative on European Language Curriculum; <i>The Threshold Level</i> (Van Ek, Council of Europe (CoE) 1975); <i>Notional Syllabuses</i> (Wilkins 1976); The notional-functional syllabus (CoE); English for Specific purposes (ESP) (Munby) 1978
<b>1975</b>	<b>The First Certificate in English (FCE)</b> <b>Dedicated Reading and Listening papers in FCE and CPE</b>
<b>1978</b>	<i>Teaching Language as Communication</i> (Widdowson 1978)
<b>1980</b>	<b>Preliminary English Test (PET); ELTS test</b>
<b>1987–9</b>	<b>The Cambridge–TOEFL Comparability Study (Bachman, Davidson, Ryan and Choi 1995)</b>
<b>1988</b>	<b>Peter Hargreaves, appointed Head of the EFL Division, arrives from the British Council</b>
<b>1989</b>	<b>IELTS test</b> <b>Creation of the EFL Evaluation Unit (later the ESOL Research and Validation Group)</b>
1990	<i>Fundamental Considerations in Language Testing</i> (Bachman 1990)
<b>1991</b>	<b>Certificate in Advanced English (CAE)</b>
<b>1993</b>	<b>Business English Certificates (BEC)</b>
<b>1994</b>	<b>Key English Test (KET)</b>
2001	<i>The Common European Framework of Reference for Languages: Learning, Teaching, Assessment</i> (CEFR) (Council of Europe 2001)
<b>2003 onwards</b>	<b>Studies in Language Testing (SiLT) ‘Constructs’ project: a socio-cognitive approach</b>

In addition, the authors looked more widely at external, institutional, social and economic forces to help us understand further the shifts in teaching and testing practice. Following World War Two (1939–45), for example, we found that it was no longer just the influence of traditional attitudes to language and language learning or the insights from linguistics or developments in modern language pedagogy that influenced examinations; increasingly, influence was exerted by the economic and socio-political forces that were at work in making English a dominant language around the globe. Traditional approaches such as grammar translation and teaching English as an access route to great literature were to succumb to pressing utilitarian needs for English as a means of communication between people rather than as a rarefied object of academic study. Changes in Cambridge English examinations in the second half of the 20th century reflected this mind shift. As Stern astutely observed, the interest in language became ‘social’ rather than ‘scholarly’ (Stern 1983:81).

Progress towards a European Economic Community from the 1970s onwards brought with it a felt need on the part of intergovernmental agencies in Europe to define language teaching and learning goals more precisely and to make a start on delineating stages in that progression. In this climate, emerging insights from research in the developing field of applied linguistics facilitated the shift on the part of the examination boards towards more explicit specification of the constructs underlying their English language tests. With the additional emphasis on communication came the need, and with developments in applied linguistics and language pedagogy perhaps the capacity, to be more explicit not just about the constructs being measured but how the measurement of these might differ according to the learner’s level of language proficiency. This saw the appearance of more Cambridge English language examinations across the proficiency spectrum from 1980 onwards (PET at the Common European Framework of Reference for languages (CEFR) B1 level in 1980, CAE at the CEFR C1 level in 1991 and KET at the CEFR A2 level in 1994). It also led to a granular approach to construct definition at different proficiency levels in Europe, which was never a major concern for testers in the United States.

At various points in *Measured Constructs* the authors compared what was happening in the United Kingdom with developments in the United States, the two world leaders in the field of English language testing in the 20th century. This was informative and helped to elucidate some of the differing, if less clear-cut, emphases that remain in approaches to theory and practice in language assessment in these two countries. Although ESOL testing in both countries is now similarly informed by all aspects of construct validity this was not always the case in the 20th century, and the path testing was to take differed markedly in each.

Substantive differences grew between the United Kingdom and the United States in their approaches to language testing from 1913–1970. An important reason for this Atlantic split could be found in the differing socio-economic climates prevailing in Britain and the United States in the early 20th century. In the United States there was a compelling need to produce tests on an industrial scale due to the population explosion in the schools system in this period, as well as a need to allocate roles to more than a million soldiers in the US army during World War One. These factors led testing organisations in the United States to adopt objective multiple-choice methods much earlier than in the United Kingdom. In the United States the predominant early focus was to be on scoring validity, in particular the psychometric qualities of a test whereas in the United Kingdom we identified a greater concern with context validity and relating examinations to what was going on in the English language teaching classroom; this could be characterised as a concern with the *how* in the United States as against the *what* in the United Kingdom. The reasons behind these contrasting journeys could often be found in the prevailing, wider socio-economic contexts but the differing approach in the United Kingdom also reflected European legacies from the past in both theoretical and practical approaches to language teaching and assessment.

Looking back over 100 years of Cambridge English examinations through the lens of CPE (see Figure 1) it is possible to discern a number of trends in its various constituent papers.

Figure 1: CPE 1913–2012

Phonetics	1913–32	
Dictation	1913–75	
Listening Comprehension		1975–
Reading aloud	1913–86	
Conversation/Oral	1913–	
Translation	1913–75 (1988)	
Essay/Composition	1913–	
Literature	1913–75	
Reading Comprehension		1975–
Knowledge of grammar	1913–32	
Use of English		1956–

Stability is one of the twin pillars of public examinations that is essential if exams are to fulfil the purposes for which they are intended. The presence of direct tests of writing and speaking throughout the history of Cambridge English examinations is testimony to this constancy, as is the length of time other components of the examination were to last (see Figure 1). Furthermore, this history shows how Cambridge was able to achieve this stability while at the same time gradually incorporating test tasks that reflected new developments in language pedagogy, linguistics and applied linguistics, first from the Grammar-Translation Method and Reform Movement approaches in Europe, then the Oral-Structural-Situational Approach and finally the Communicative Approach. This second pillar, innovation linked to improvement, is just as vital if an examination is to keep up with developments and insights available from research in the field. Thus in the 1950s we see a Use of English paper making an appearance with a number of tasks reflecting the increasing interest in the structural approach to language teaching. By the 1970s the early influence of the academic, scholarly view of language, which regarded it merely as an object of study, had largely disappeared with increased attention being paid to the way language was used for communication. Gone was the prominence given to a knowledge of phonetics, translation, English literature, and grammatical usage to be largely replaced by papers on listening and reading skills in their own right, plus a revamped Speaking test. We also see the gradual introduction of new technology into the Listening test in the 1980s as attempts were made to make both texts and tasks more 'authentic' in line with developments in communicative language teaching. The more traditional integrated tests of literature, translation, dictation and reading aloud had disappeared by the 1980s just as phonetics and grammar had in the 1930s.

Figure 1 also shows how Cambridge responded to the socio-economic impact of events in the wider world, in particular to the globalisation of English that gathered momentum after World War Two. The growth of English as a world language was reflected in the introduction of a language-only route for CPE examination candidates and the downgrading of the importance of literature and cultural knowledge in the overall language ability construct. The idea of English as an international language was also reflected in the later development of the speaking assessment scales in the 1980s and the downgrading of the native speaker concept as the top of the scale.

This historical survey suggests that *context validity* had been a major focus of attention for the Cambridge English tests throughout most of the 20th century. There was to be additional improvement to the validity of test scores at the University of Cambridge Local Examinations Syndicate (UCLES) EFL in the late 1980s when, largely as a result of the Cambridge-TOEFL Comparability Study (Bachman, Davidson, Ryan and Choi 1995), far greater attention was paid to *scoring validity* to bring Cambridge's procedures in line with the psychometrically more sophisticated approach that had long been part of professional language assessment in the United States. The change process was facilitated by the appointment of a professional core of research staff receptive to such ideas at the heart of the organisation's activities, starting with the

arrival of Peter Hargreaves in 1988, closely followed by Neil Jones, Michael Milanovic, Nick Saville and Lynda Taylor. Thus *scoring validity* finally established its place as a fundamental canon of the examination system in Cambridge by the end of the 20th century. In the early 21st century, the focus extended to include *cognitive validity* as a result of producing the 'construct' volumes, over a 10-year period, in the SiLT series, guided by Michael Milanovic, Nick Saville, Lynda Taylor and Cyril Weir on the editorial steering committee. This ambitious project enabled far greater attention to be paid than previously to the cognitive processing typically activated in test and non-test tasks, and to the importance of an appropriate match between the two. There is now a growing recognition within Cambridge English Language Assessment itself and its partners, and in the wider international testing community, of the importance for any successful assessment system of seeking and assembling validity evidence on each of these three core aspects of validity: cognitive, context and scoring, which together constitute test construct validity.

An overt concern with the constructs being measured in the Cambridge English examinations and their relationship to real-life language use was apparent by the end of the historical survey. This commitment to transparency and the explicit specification of the communicative content of its examinations was enhanced by Cambridge's interest in a socio-cognitive approach to language test design and validation in the first decade of the 21st century; such an approach acknowledges that language use constitutes both a socially situated and a cognitively processed phenomenon and that this must be reflected in language assessment theory and practice.

## Endnote

*Measured Constructs* (Weir et al forthcoming 2013) sheds light on how approaches to measuring language constructs evolved at Cambridge in the 100 years of its English language examinations. It takes the reader from the first form of CPE (*Cambridge English: Proficiency*) offered to three candidates in 1913, a serendipitous hybrid of legacies in language teaching from the previous century, up to the 2012 Cambridge approach to language examinations, where the language construct to be measured in the test is seen as an evidence-based product of the interaction between a targeted cognitive ability based on an expert user model, a highly specified context of use and a performance level based on explicit and appropriate criteria of description.

## References

- Bachman, L F, Davidson, F, Ryan, K and Choi, I C (1995) *An Investigation into the Comparability of Two Tests of English as a Foreign Language*, Studies in Language Testing volume 1, Cambridge: UCLES/Cambridge University Press.
- Bloomfield, L (1926) A set of postulates for the science of language, *Language* 2, 153-164.
- Bloomfield, L (1933) *Language*, New York: Henry Holt.

- Brooks, N (1960) *Language and Language Learning: Theory and Practice*, New York: Harcourt, Brace.
- Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.
- Eckersley, C E (1938–42) *Essential English for Foreign Students*, volumes 1–4, London: Longmans, Green.
- Edgeworth, F Y (1888) The Statistics of Examinations, *Journal of the Royal Statistical Society*, 599–635.
- Faucett, L W (1933–34) *The Oxford English Course*, volumes 1–4, London: Oxford University Press.
- Faucett, L W, Palmer, H E, Thorndike, E L and West, M P (1936) *Interim Report on Vocabulary Selection for the Teaching of English as a Foreign Language*, London: King.
- Fick, J C (1793) *Praktische englische Sprachlehre für Deutsche beyderley Geschlechts. Nach der in Meidingers französichen Grammatik befolgten Methode* [Practical English Language Textbook for Germans of Both Sexes. Following the Method Used in Meidinger's French Grammar], Erlangen.
- Fries, C C (1945) *Teaching and Learning English as a Foreign Language*, Michigan: University of Michigan Press.
- Fries, C C and Lado, R (1962) *English Sentence Patterns, Understanding and Producing Grammatical Structures*, Ann Arbor: University of Michigan Press.
- Geranpayeh, A and Taylor, L (Eds) (forthcoming 2013) *Examining Listening: Research and Practice in Assessing Second Language Listening*, Studies in Language Testing volume 35, Cambridge: UCLES/Cambridge University Press.
- Hornby, A S (1950) The situational approach in language teaching, *ELT Journal*, 4–6.
- Hornby, A S (1954–56) *Oxford Progressive English for Adult Learners*, London: Oxford University Press.
- Howatt, A P R (1984) *A History of English Language Teaching*, Oxford: Oxford University Press.
- Howatt, A P R and Widdowson, H G (2004) *A History of English Language Teaching* (Second edition), Oxford: Oxford University Press.
- Jespersen, O (1904) *How to Teach a Foreign Language*, London: G Allen and Unwin.
- Khalifa, H and Weir, C J (2009) *Examining Reading: Research and Practice in Assessing Second Language Reading*, Studies in Language Testing volume 29, Cambridge: UCLES/Cambridge University Press.
- Meidinger, J V (1783) *Praktische französische Grammatik* [Practical French Grammar], Frankfurt.
- Munby, J L (1978) *Communicative Syllabus Design*, Cambridge: Cambridge University Press.
- Palmer, H E (1921a) *The Oral Method of Teaching Languages*, Cambridge: Heffer and Sons Ltd.
- Palmer, H E (1921b) *The Principles of Language-Study*, London: Harrap.
- Palmer, H E and Redman, H V (1932) *This Language Learning Business*, London: Harrap.
- Passy, P (1899) *De la méthode directe dans l'enseignement des langues vivantes* [The Direct Method in the Teaching of Modern Languages], Paris: Colin.
- Richards, J C and Rogers, T S (2001) *Approaches and Methods in Language Teaching* (Second edition), Cambridge: Cambridge University Press.
- Shaw, S D and Weir, C J (2007) *Examining Writing: Research and Practice in Assessing Second Language Writing*, Studies in Language Testing volume 26, Cambridge: UCLES/Cambridge University Press.
- Shohamy, E (2008) Introduction, in Shohamy, E and Hornberger, N (Eds) *Encyclopedia of Language and Education* (Second edition), Language Testing and Assessment, volume 7, New York: Springer Science+Business Media LLC, xiii–xxii.
- Spolsky, B (1995) *Measured Words*, Oxford: Oxford University Press.
- Spolsky, B (2008) Introduction – Language Testing at 25: Maturity and responsibility? *Language Testing* 25 (3), 297–305.
- Stern, H H (1983) *Fundamental Concepts of Language Teaching*, Oxford: Oxford University Press.
- Sweet, H (1899) *The Practical Study of Languages*, London: Dent.
- Taylor, L (2009) Developing assessment literacy, *Annual Review of Applied Linguistics* 29, 21–36.
- Taylor, L (Ed.) (2011) *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, Studies in Language Testing volume 30, Cambridge: UCLES/Cambridge University Press.
- University of Cambridge Local Examinations Syndicate (1913) *Regulations for the Examinations for Certificates of Proficiency in Modern Languages and Religious Knowledge*, Cambridge: University of Cambridge Local Examinations Syndicate.
- University of Cambridge Local Examinations Syndicate (1947) *Regulations*, Cambridge: University of Cambridge Local Examinations Syndicate.
- Van Ek, J (1975) *Systems Development in Adult Language Learning: The Threshold Level*, Strasbourg: Council of Europe.
- Viëtor, W (1882) *Der Sprachunterricht muss umkehren! Ein Beitrag zur Überbürdungsfrage* [Language teaching must turn round! A Contribution to the Overloading Question], (Under the pseudonym Quousque Tandem), Heilbronn: Henninger.
- Weir, C J and Milanovic, M (Eds) (2003) *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913–2002*, Studies in Language Testing volume 15, Cambridge: UCLES/Cambridge University Press.
- Weir, C J, Vidaković, I and Galaczi, E D (forthcoming 2013) *Measured Constructs: A History of Cambridge English Language Examinations 1913–2012*, Studies in Language Testing volume 37, Cambridge: UCLES/Cambridge University Press.
- Widdowson, H G (1978) *Teaching Language as Communication*, Oxford: Oxford University Press.
- Widdowson, H G (2003) *Defining Issues in English Language Teaching*, Oxford: Oxford University Press.
- Wilkins, D A (1976) *Notional Syllabuses*, London: Oxford University Press.
- Wyatt, T S and Roach, J O (1947) The examinations in English of the Cambridge University Local Syndicate, *ELT Journal* 1 (5), 125–130.

## Appendix: 1913 Certificate of Proficiency in English exam

<p style="text-align: center;"><b>Certificate of Proficiency in English</b></p> <p style="text-align: center;"><b>Translation into French</b></p> <p style="text-align: center;">(Two hours.)</p> <p>My grandfather's recollections of Culloden were merely those of an observant boy of fourteen, who had witnessed the battle from a distance. The day, he has told me, was drizzly and thick; and on reaching the brow of the Hill of Cromarty, where he found many of his townfolk already assembled, he could scarce see the opposite land. But the fog gradually cleared away; first one hill-top came into view, and then another; till at length the long range of coast, from the opening of the great Caledonian valley to the promontory of Burghead, was dimly visible through the haze. A little after noon there suddenly rose a round white cloud from the Moor of Culloden, and then a second round white cloud beside it. And then the two clouds mingled together, and went rolling slantways on the wind towards the west; and he could hear the rattle of the smaller fire-arms mingling with the roar of the artillery. And then, in what seemed an exceedingly brief</p>	<p style="text-align: center;"><b>Certificate of Proficiency in English</b> 2</p> <p>space of time, the cloud dissipated and disappeared, the boom of the greater guns ceased, and a sharp intermittent patter of musketry passed on towards Inverness. But the battle was presented to the imagination, in these old personal narratives, in many a diverse form. I have been told by an ancient woman, who, on the day of the fight, was engaged in tending some sheep on a solitary common near Munloch, separated from the Moor of Culloden by the Frith, and screened by a lofty hill, that she sat listening in terror to the boom of the cannon; but that she was even still more scared by the continuous howling of her dog, who sat upright on his haunches all the time the firing lasted, with his neck stretched out towards the battle, and "looking as if he saw a spirit." Such are some of the recollections which link the memories of a man who has lived his half-century to those of the preceding age, and which serve to remind him how one generation of men after another break and disappear on the shores of the eternal world, as wave after wave breaks in foam upon the beach, when storms are rising, and the ground-swell sets in heavily from the sea.</p> <p style="text-align: right;">HUGH MILLER.</p> <p style="text-align: center;"><b>Translation into German</b></p> <p style="text-align: center;">(Two hours.)</p> <p>Translate into GERMAN:</p> <p>(a) The sentiments which animated Schiller's poetry were converted into principles of conduct; his actions were as blameless as his writings were pure. With his simple and high predilections, with his strong devotedness to a noble cause, he contrived to steer through life, unsullied by its meanness, unsubdued by any of its difficulties or allurements. With the world, in fact, he had not much to do; without effort he dwelt apart from it; its prizes were not the wealth which could enrich him. His great, almost his single aim, was to unfold his spiritual faculties, to study and contemplate and</p>
<p>3</p> <p style="text-align: center;"><b>Certificate of Proficiency in English</b></p> <p>improve their intellectual creations. Bent upon this, with the steadfastness of an apostle, the more sordid temptations of the world passed harmlessly over him. Wishing not to seem, but to be, envy was a feeling of which he knew but little, even before he rose above its level. Wealth or rank he regarded as means, not an end; his own humble fortune supplying him with all the essential conveniences of life, the world had nothing more that he chose to covet, nothing more that it could give him. In fact his real wealth lay in being able to pursue his darling studies, and to live in the sunshine of friendship and domestic love. This he had always longed for; this he at last enjoyed.</p> <p style="text-align: right;">THOMAS CARLYLE.</p> <p>(b) Early this morning we left Aix-la-Chapelle and came on to Cologne. The country, which about Aix is very pretty, soon degenerates into great masses of table-land which at last sank down into a plain, and from the edge of it, as we began to descend, we burst upon the view of the valley of the Rhine, the city of Cologne, with all its towers, the Rhine itself distinctly seen at the distance of seven miles—the Seven Mountains above Bonn on our right, and a boundless sweep of the country beyond the Rhine in front of us. To be sure, it was a striking contrast to the first view of the valley of the Tiber from the mountain of Viterbo; but the Rhine in mighty recollections will vie with anything, and this spot was particularly striking: Cologne was Agrippa's colony (<i>Colonia Agrippina</i>), inhabited by Germans brought from beyond the river, to live as the subjects of Rome; the river itself was the frontier of the Empire—the limit as it were of two worlds, that of Roman laws and customs and that of German. Far before us lay the land of our Saxon and Teutonic forefathers—the birthplace of the most moral races of men that the world has yet seen—of the soundest laws—the least violent passions, and the fairest civic and domestic virtues. I thought of that memorable defeat of Varus and his three legions which for ever confined the Romans to the Western side of the Rhine, and preserved the Teutonic nation—the regenerating element in modern Europe—safe and free.</p> <p style="text-align: right;">THOMAS ARNOLD.</p>	<p style="text-align: center;"><b>Certificate of Proficiency in English</b> 4</p> <p style="text-align: center;"><b>Translation into English and English Grammar</b></p> <p style="text-align: center;">(Two and a half hours.)</p> <p style="text-align: center;">PART I. TRANSLATION</p> <p style="text-align: center;">A. (For candidates taking French.)</p> <p>Translate into ENGLISH:</p> <p>(a) On entre d'abord dans un corridor large et bien éclairé, mais dont la largeur est diminuée par de vastes armoires de noyer sculpté, où les paysans enferment le linge du ménage, et par des sacs de blé ou de farine, déposés là pour les besoins journaliers de la famille. A gauche est la cuisine, dont la porte, toujours ouverte, laisse apercevoir une longue table de bois de chêne entourée de bancs. Il est rare qu'on n'y voie pas des paysans attablés à toute heure de jour; car la nappe y est toujours mise, soit pour les ouvriers, soit pour ces innombrables survenants à qui on offre habituellement le pain, le vin et le fromage, dans des campagnes éloignées des villes et qui n'ont ni auberge ni cabaret. A droite on entre dans la salle à manger. Rien ne la décore qu'une table de sapin, quelques chaises et un de ces vieux buffets à compartiments, à tiroirs et à nombreuses étagères, meuble héréditaire dans toutes les vieilles demeures, et que le goût actuel vient de rajeunir en les recherchant. De la salle à manger on passe dans un salon à deux fenêtres, l'une sur la cour, l'autre au nord, sur un jardin. Un escalier, en bois, mène à l'étage unique et bas où une dizaine de chambres presque sans meubles ouvrent sur des corridors obscurs. Elles servaient alors à la famille, aux hôtes et aux domestiques. Voilà tout l'intérieur de cette maison qui nous a si longtemps couvés dans ses murs sombres et chauds; voilà le toit que ma mère appelait avec tant d'amour sa "Jérusalem," sa maison de paix; voilà le nid qui nous abrita tant d'années de la pluie, du froid, de la faim, du souffle du monde.</p> <p style="text-align: right;">LAMARTINE.</p>

5	Certificate of Proficiency in English	Certificate of Proficiency in English	6
	<p>(b) Nous venons de faire un empereur, et pour ma part je n'y ai pas nui. Voici l'histoire. Ce matin, d'Anthonard nous assemble, et nous dit de quoi il s'agissait, mais bonnement, sans préambule ni péroraison. Un empereur ou la république, lequel est le plus de votre goût? comme on dit, Rôti ou bouilli, potage ou soupe, que voulez-vous? Sa harangue finie, nous voilà tous à nous regarder, assis en rond. Messieurs, qu'opinez-vous? Pas le mot; personne n'ouvre la bouche. Cela dura un quart d'heure ou plus, et devenait embarrassant pour d'Anthonard et pour tout le monde, quand Maire, un jeune homme un lieutenant que tu as pu voir, se lève, et dit: S'il veut être empereur, qu'il le soit; mais, pour en dire mon avis, je ne le trouve pas bon du tout. Expliquez-vous, dit le colonel; voulez-vous? ne voulez-vous pas? Je ne le veux pas, répond Maire. A la bonne heure. Nouveau silence. On recommence à s'observer les uns les autres, comme des gens qui se voient pour la première fois. Nous y serions encore, si je n'eusse pris la parole. Messieurs, dis-je, il me semble, sauf correction, que ceci ne nous regarde pas. La nation veut un empereur, est-ce à nous d'en délibérer? Ce raisonnement parut si fort, si lumineux, si <i>ad rem</i>...que veux-tu? j'entraînai l'assemblée. Jamais orateur n'eut un succès si complet. On se lève, on signe, on s'en va jouer au billard. Maire me disait: Ma foi, commandant, vous parlez comme Cicéron; mais pourquoi voulez-vous donc tant qu'il soit empereur, je vous prie? Pour en finir, et faire notre partie de billard. Fallait-il rester là tout le jour? pourquoi, vous, ne le voulez-vous pas? Je ne sais, me dit-il, mais je le croyais fait pour quelque chose de mieux. Voilà le propos du lieutenant, que je ne trouve point tant sot. En effet, que signifie, dis-moi...un homme comme lui, Bonaparte, soldat, chef d'armée, le premier capitaine du monde, vouloir qu'on l'appelle Majesté? Etre Bonaparte, et se faire sire! <i>Il aspire à descendre</i>: mais non, il croit monter en s'égalant aux rois. Il aime mieux un titre qu'un nom. Pauvre homme! ses idées sont au-dessous de sa fortune. Je m'en doutai quand je le vis donner sa petite sœur à Borghèse, et croire que Borghèse lui faisait trop d'honneur.</p> <p style="text-align: right;">P. L. COURIER.</p>	<p>B. (For candidates taking German.)</p> <p>Translate into ENGLISH:</p> <p>(a) Aber der augenfälligste Zug im gesammten Leben des Harzes ist das berg- und hüttenmännische Wesen. Das haben seine Blei- und Silbergruben, vorzüglich der Gegenden von Goslar, Klausthal, Zellerfeld, Andreasberg und Harzgerode, veranlaßt. Daher findet sich dort eine große Zahl Anstalten, die sich auf den Bergbau beziehen; daher nehmen dort nicht bloß dampfende Kohlenmeiler, Waldarbeiten aller Art und die rein und harmonisch tönenden Glocken der Viehherden, mit denen die einsamen Hirten weit in die Wälder hineinziehen, sondern weit mehr noch die bunten Erscheinungen bergmännischen Treibens unsere Aufmerksamkeit und Teilnahme in Anspruch. Denn überall schwingt dort der Bergmann den Häufel, schmilzt der braune Hüttenmann die dem Gebirgsschooße entnommenen Erze; überall sieht man dort Gruben, aufsteigende Rauchwolken, Karren mit Erz in unaufhörlicher Bewegung, und in frühlicher Stunde vernimmt man in der harten, scharfen Mundart den Trinkspruch des Oberharzes: „Es grüne die Tanne, es wachse das Erz! Gott schenke uns allen ein frühliches Herz!“ Und so zeigt sich das Harzleben seit einer Reihe von Jahrhunderten. Schon in der zweiten Hälfte des zehnten Jahrhunderts sollen die dortigen Silberbergwerke entdeckt worden sein; mit Gewißheit befinden im elften Jahrhunderte geschichtliche Zeugnisse eine Fülle des edeln Metalls als Mittel zum Verkehr und zu Kunstgegenständen. Wie würden ohne den Erzreichtum des Gebirges von der Bevölkerung so früh gar manche an sich wenig ansehende Höhen zu Ansiedlungen gewährt werden, oder wie würden noch in unserem Zeitalter auf ihnen in Gegenden gastliche Stätten entstanden sein, wo sonst kaum ärmliche Hütten zu finden waren und wo ringsum auch heute weder Obstbäume noch Saatzfelder gesehen, nur einige Kuckengewächse dürrig erzieht werden?</p> <p style="text-align: right;">J. KUTZEN.</p> <p>(b) Besonders machen die Quacqueri zwar nicht so viel Lärm, doch eben so viel Aufsehen als die Abvofaten. Die Maske der Quacqueri scheint so allgemein geworden zu sein durch die Leichtigkeit, auf dem Trödel altfränkische Kleidungsstücke finden zu können. Die Haupterfordernisse dieser Maske sind, daß die Kleidung zwar altfränkisch, aber wohlhalten und von edlem Stoff sei. Man sieht sie</p>	

7	Certificate of Proficiency in English	Certificate of Proficiency in English	8
	<p>seifen anders, als mit Sammt oder Seide bekleidet, sie tragen brofatene oder gestifte Westen, und der Natur nach muß der Quacquero dieleibig sein; seine Gesichtsmaske ist ganz, mit Bausbaden und kleinen Augen; seine Perrücke hat wunderliche Zöpfchen; sein Hut ist klein und meistens bordiert. Man sieht, daß sich diese Figur sehr dem Buffo caricato der komischen Oper nähert, und wie dieser meistens einen läppischen, verlichten, betrogenen Loren vorstellt, so zeigen sich auch diese als abgeschmackte Stutzer. Sie hüpfen mit großer Leichtigkeit auf den Beinen hin und her, führen große schwarze Ringe ohne Glas statt der Vergnieten, womit sie in alle Wagen hineingucken, nach allen Fenstern hinaufblicken. Sie machen gewöhnlich einen steifen, tiefen Büdfling, und ihre Freude, besonders wenn sie sich einander begegnen, geben sie dadurch zu erkennen, daß sie mit gleichen Füßen mehrmals gerade in die Höhe hüpfen und einen hellen durchdringenden unartikulierten Laut von sich geben, der mit den Konsonanten <i>ber</i> verbunden ist. Oft geben sie sich durch diesen Ton das Zeichen, und die nächsten erwidern das Signal, so daß in kurzer Zeit dieses Geschreie den ganzen Corso hin und wieder läuft. Mutwillige Knaben blasen indeß in große gewundene Muscheln und beleidigen das Ohr mit unerträglichen Tönen.</p> <p style="text-align: right;">GOETHE.</p> <p style="text-align: center;">PART II. GRAMMAR</p> <p style="text-align: center;">(For all candidates.)</p> <p>1. Give the past tense and past participle of each of the following verbs, dividing them into strong and weak; add explanations: <i>tell, wake, buy, eat, lay, lie</i>.</p> <p>2. Mark by an acute accent the accented syllable in each of the following words: <i>subjected, hyperbole, microscopical, photography, contemplative, confident, confidant, pusillanimity, gangrene, tureen</i>.</p> <p>3. Write down the abstract nouns connected with the following adjectives and verbs: <i>precise, adhere, apt, predominate, optimistic, crystallise, negligent, hate, attain, detain, betray, ingenious, seize, charitable, zealous</i>.</p>	<p>4. Embody each of the following words in a sentence, in such a way as to shew that you clearly apprehend its meaning: <i>commence, comment, commend, recommend; incredible, incredulous</i>.</p> <p>5. Correct or justify <b>four</b> of the following sentences, giving your reasons:</p> <p>(a) I hope you are determined to seriously improve.</p> <p>(b) Comparing Shakespeare with Æschylus, the former is by no means inferior to the latter.</p> <p>(c) I admit that I was willing to have made peace with you.</p> <p>(d) The statement was incorrect, as any one familiar with the spot, and who was acquainted with the facts, will admit.</p> <p>(e) It has the largest circulation of any paper in England.</p> <p>(f) The lyrical gifts of Shakespeare are woven into the actual language of the characters.</p> <p style="text-align: center;">English Essay</p> <p style="text-align: center;">(Two hours.)</p> <p>Write an Essay on <b>one</b> of the following subjects:</p> <p>(a) The effect of political movements upon nineteenth century literature in England.</p> <p>(b) English Pre-Raffaellism.</p> <p>(c) Elizabethan travel and discovery.</p> <p>(d) The Indian Mutiny.</p> <p>(e) The development of local self-government.</p> <p>(f) Matthew Arnold.</p>	

<p>9</p> <p><b>Certificate of Proficiency in English</b></p> <p><b>English Phonetics</b></p> <p><i>(One hour and a half.)</i></p> <p><i>Candidates may use any consistent system of phonetic notation, but must state what system they use; and if the system they use is not generally known, they should shew by examples what values are to be attached to the symbols they use.</i></p> <p>1. Make a phonetic transcription of each of the following passages, illustrating in the case of passage (a) a careful pronunciation, in the case of (b) the pronunciation of educated persons in ordinary conversation :</p> <p>(a) But, whatever be the profession or trade chosen, the advantages are many and important, compared with the state of a mere literary man, who in any degree depends on the sale of his works for the necessaries and comforts of life. In the former a man lives in sympathy with the world in which he lives. At least he acquires a better and quicker tact for the knowledge of that with which men in general can sympathise. He learns to manage his genius more prudently and efficaciously. His powers and acquirements gain him likewise more real admiration; for they surpass the legitimate expectations of others. He is something besides an author, and is not therefore considered merely as an author. The hearts of men are open to him as to one of their own class; and whether he exerts himself or not in the conversational circles of his acquaintance, his silence is not attributed to pride, nor his communicativeness to vanity.</p> <p>(b) "Ah, Mr Holmes. I am delighted to see you."  "Good morning, Lanner. You will not think me an intruder, I am sure. Have you heard of the events which led up to this affair?"  "Yes, I heard something of them."  "Have you formed any opinion?"  "As far as I can see, the man has been driven out of his senses by fright. The bed has been well slept in, you see. There is his impression deep enough."</p>	<p><b>Certificate of Proficiency in English</b></p> <p>10</p> <p>"Noticed anything peculiar about the room?"  "Found a screwdriver and some screws on the wash-hand stand. Seems to have smoked heavily during the night, too. Here are four cigar ends that I picked out of the fire-place."  "Hum! Have you got his cigar-holder?"  "No, I have seen none."  "His cigar case then?"  "Yes, it was in his coat pocket."</p> <p>2. Describe fully the articulation of the various vowel sounds in the (ordinary) spelling of which the letter <i>o</i> is used (alone or in combination) in the above passages.</p> <p>3. Explain the terms: 'glide,' 'narrow vowel,' 'semi-vowel,' and give two examples of each in both phonetic and ordinary spelling.</p> <p>4. How would you teach a pupil the correct pronunciation of the vowel sounds in <i>fare, fate, fat, fall, far</i>?</p> <p>5. Discuss carefully the articulation of the consonants in <i>quite, huge, dreary</i>.</p>
--	---

# Components of an elaborated approach to test validation

GAD S LIM RESEARCH AND VALIDATION GROUP, CAMBRIDGE ENGLISH LANGUAGE ASSESSMENT

## Introduction

Among the things that have seen further development and elaboration over the last decade is the Cambridge English Language Assessment approach to test validation. In the volume which detailed the revision process for the 2002 version of *Cambridge English: Proficiency*, Saville (2003) wrote about fundamental considerations in developing Cambridge English examinations. The considerations in developing useful tests, informed in part by seminars conducted by Professor Lyle Bachman in Cambridge in 1991, include validity, reliability, impact and practicality (VRIP).

The validation of tests is intrinsically something that happens over time, and it is important to account for this temporal dimension of testing as well. In the same chapter, Saville (2003) stressed the cyclical and iterative nature of test development, and briefly mentioned initial steps to engage with ideas regarding quality management systems that are employed in manufacturing and service industries. The relationship between quality management and VRIP was, however, not made explicit then.

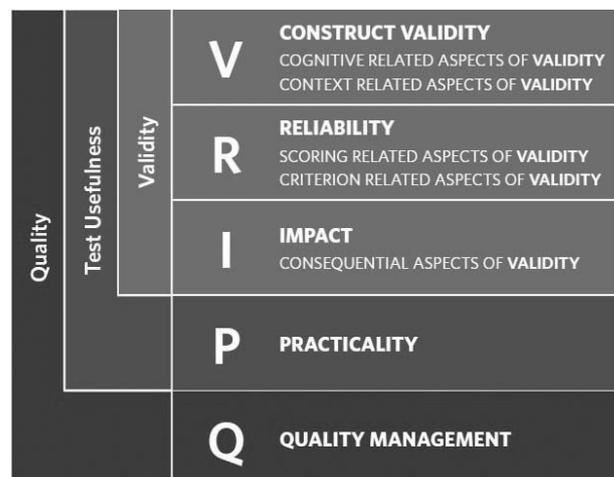
Another development in the last decade has been the adoption of a socio-cognitive framework for test validation, which was developed in collaboration with and detailed in Weir (2005). The framework lays out the various aspects of validity to account for in test validation and their relationship with each other. These include cognitive-related, context-related, scoring-related, criterion-related, and consequential aspects of validity. Again, the relationship between this framework and VRIP may not be immediately transparent. For example, validity is a consideration in both, so there is a seeming overlap between them.

In view of the above, this article discusses how the developments over the last decade relate to VRIP. In so doing, it presents a picture of the elaborated approach to test validation that is employed by Cambridge English Language Assessment at present. A brief discussion of the 2013 revision to *Cambridge English: Proficiency* is also included to illustrate how the elaborated approach has informed actual test validation work.

## An elaborated approach to test validation

The relationship between the various components mentioned above is perhaps made clearer when illustrated. Figure 1 shows VRIP, quality management, and the socio-cognitive framework's aspects of validity in horizontal boxes on the right side. In addition, the figure also includes a sidebar on the left, the vertical text capturing overall features of tests. These are further explained in the next sections.

**Figure 1: The elaborated Cambridge English Language Assessment approach to test validation (Cambridge English Language Assessment 2013)**



### VRIP, quality management, and the socio-cognitive framework

The original formulation of VRIP captured the different aspects that needed to be accounted for in creating tests that are valid and useful, and in that sense Cambridge English Language Assessment had always considered all of them in developing and revising tests. Examples of this can be found in Weir and Milanovic (Eds) (2003) and Hawkey (2009).

It is possible, however, for VRIP to be considered only during initial test validation whereas, as has been previously mentioned, test validation is a continuing activity. For example, a situation can be imagined where a good prototype test form is not followed by comparable/equally good test forms, perhaps because the attention and resources lavished on the test at first was not maintained subsequently. This example makes clear that test validation, as a repeated activity that happens over time, benefits from having defined processes, e.g. for how things are done, for checking that these processes are being followed. This is the function that quality management fulfills, and why the Cambridge English Language Assessment approach to test validation has been extended from VRIP to VRIPQ.

VRIP and quality management can be seen as two sides of the same coin. VRIP captures the *what* of test validation, and quality management captures the *how* of managing the necessary processes of test validation. That is, it 'provides a practical approach to putting principles into practice' (Saville 2012b:409). An assessment provider with quality systems in place is, all things being equal, more likely to end up with a more comprehensively validated test.

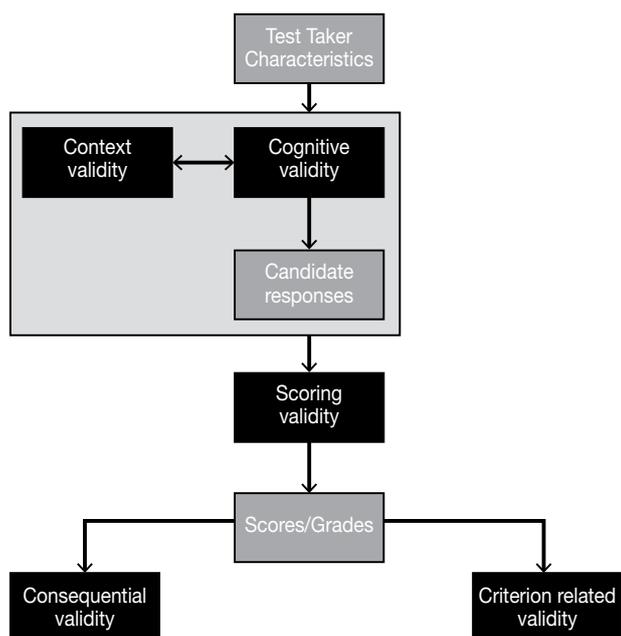
In practice, quality activities include those having to do with quality control (checking a product to make sure it meets requirements) and those having to do with quality assurance (managing and monitoring production processes).

The Cambridge English Language Assessment quality management systems are certified to ISO 9001, comply with standards set out by the UK government regulator Ofqual, and are regularly evaluated by internal and external auditors. For further examples of Cambridge English Language Assessment quality activities, see *Principles of Good Practice: Quality Management and Validation in Language Assessment* (Cambridge English Language Assessment 2013).

The socio-cognitive framework for test validation can be seen as an elaboration of the different aspects of a valid test so that these different aspects might be properly accounted for and validated in a structured and systematic way. The different aspects of the socio-cognitive framework are thus subsumed under the V, R, and I boxes of VRIP in Figure 1. The framework makes it clear, for example, that to properly address the issue of reliability, one needs to account for both scoring-related and criterion-related aspects of it. Weir (2005) further lists the specifics of each of those aspects, providing the practitioner with detailed guidance on validating language tests.

In addition, the socio-cognitive framework shows the relationship between the different aspects of validity (Figure 2). While all aspects need to be considered during test development, the placement of the different aspects and the direction of the arrows make clear that construct validity – cognitive-related and context-related aspects of validity – come first, whereas other aspects of validity (e.g. consequential aspects) can only be properly and fully considered later on. In this, too, the language test-validation practitioner is given direction on the relative ordering of validation activities. Use of the socio-cognitive framework for test validation is demonstrated in four *Studies in Language Testing* volumes, dealing with reading (Khalifa and Weir 2009), writing (Shaw and Weir 2007), listening (Geranpayeh and Taylor (Eds) forthcoming 2013) and speaking (Taylor (Ed) 2011).

Figure 2: Socio-cognitive framework (Weir 2005)



### Validity, test usefulness, and quality

The sidebar in Figure 1 shows the overall qualities of tests that VRIPQ work contributes towards, namely: validity, test usefulness, and quality.

Validity, following Messick (1989:13), is understood to be ‘an integrated evaluative judgement of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores’. That validity is ‘integrated’ means that there are different aspects to it: construct validity (‘empirical evidence and theoretical rationales’), reliability (‘test scores’) and impact (‘inferences and actions based on test scores’). The socio-cognitive framework is an alternative, more detailed formulation of the different aspects of validity. That validity is ‘integrated’ also means that the different aspects cannot be evaluated independently of each other. Rather, the different aspects are considered together to determine the extent to which a test is fit for purpose. For the above reasons, *validity* covers V, R, and I in Figure 1.

It is interesting, if unsurprising, that the theoretical literature has focused on validity. For purposes of theory and research, that is the natural thing to focus on. However, for examination bodies such as Cambridge English Language Assessment, which need to deal with the realities of designing and administering tests, issues of practicality are also an important focus. Only to a limited extent has this emerged in the literature, for example in Bachman and Palmer’s (1996) concept of ‘test usefulness’. A test that is highly valid and reliable for a particular use may ultimately prove to be not very useful if it is not possible to produce multiple versions of the same, difficult to administer or extremely expensive. And a test that cannot be reproduced or administered consistently cannot be a quality test, and ultimately cannot be a valid test. Indeed, inherent in Kane’s (2006) recasting of validity as validation, from a noun into a verb, is the ongoing nature of the activity, and subscribing to this definition of validity requires consideration of test use and test processes. *Test usefulness* and *quality* (in the sidebar of Figure 1) are thus important and equal considerations in the initial and continuing validation of tests.

To sum up, tests need to be valid, useful, and of consistently high quality. A number of things need to be accounted for to ensure that that is the case, and the elaborated approach to test validation helps one to see what those different components are. VRIP communicates the essential qualities to consider in developing and revising tests. The socio-cognitive framework provides a principled and orderly approach to investigating those qualities in a test. Quality management (Q) captures the necessity of systems for the continual validation and improvement of tests over time.

### The elaborated approach and the revision of Cambridge English: Proficiency

The elaboration of the approach to test validation happened over the course of the last decade, and *Cambridge English: Proficiency* for 2013 is the first test revision/updates project since. The revision itself is in fact a product of VRIPQ, because one result of Q within Cambridge English Language

Assessment is the institution of a process whereby tests are reviewed at regular intervals, occasionally resulting in the decision to update a test product, which was what happened in this case.

VRIPQ also helped those involved in the project to see the big picture. For example, offering a computer-based version of the exam involves practicalities - which create both possibilities and limitations - that needed to be weighed alongside other considerations. Introducing a new task type, as ultimately was the case with Part 1 of the Writing paper (Spillett 2012), involves judging the sustainability of the task type; that is, whether quality can be ensured over time. In keeping with the desire to create positive 'impact by design' (Saville 2012a), work was also undertaken to consider what the updated test should be like and what effects it might have on different users (e.g. Docherty and Howden, this issue).

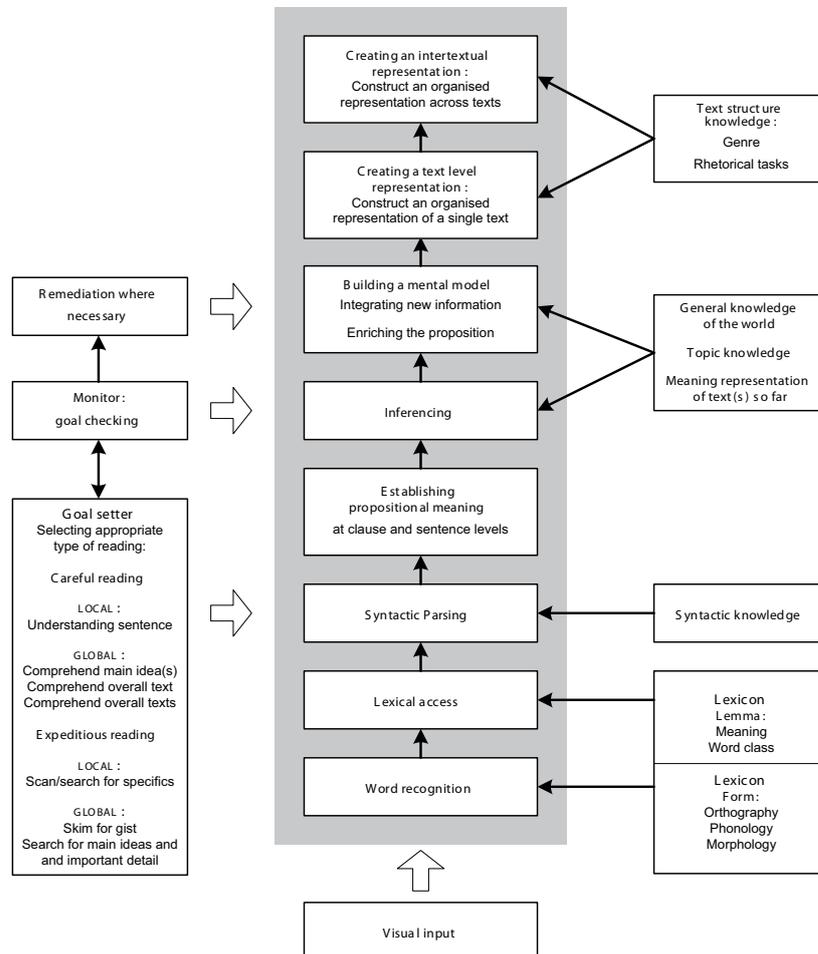
Validation work was then guided by the more detailed socio-cognitive framework. The socio-cognitive framework makes clear that during the development stage, while keeping in mind all aspects of validity, the primary focus is on construct validity (cognitive and contextual aspects) and scoring validity. Figure 2 captures this aspect of the framework by putting these aspects of validity in a larger box. Further, the framework makes clear that cognitive and contextual considerations inform each other, as indicated by the double headed arrow between them.

Cognitive aspects were already considered, in the case of reading for example, in Khalifa and Weir (2009), where a model was constructed of what happens during the reading process (from word recognition to creating representations of texts) (Figure 3). Tasks and questions were then evaluated to make sure that each level of the reading process was covered in the Reading paper. Similar analyses had been or were also conducted for the other skills (Field 2011, forthcoming 2013, Lim 2010, Zeronis and Elliott, this issue).

In addition, contextual parameters were also investigated using more quantitative methods. Structural equation modelling was used to evaluate which task types were related to each other, in order to provide evidence that tasks hypothesised to tap the same skills were indeed related to each other (Somers, Geranpayeh and Malarkey 2010). Scoring validity was also considered, for example, in the modelling of reliability, conditional standard errors of measurement, and classification consistency given various combinations of items and task types.

The different analyses above all influenced the ultimate design, selection and ordering of tasks in the new *Cambridge English: Proficiency* papers so that, while shorter, continued to cover all aspects of the construct in a reliable way. It can thus be seen how, guided by a socio-cognitive framework, a combination of theory, qualitative and quantitative methods

Figure 3: Model of the reading process (Khalifa and Weir 2009:43)



were used to consider different aspects of validity together in order to shape a test that is valid and useful.

## Conclusion

As with its other activities, the Cambridge English Language Assessment approach to test validation has exhibited continuity and innovation. Over the last decade, VRIP has been built upon and further developed. The result is an approach to test validation that is more complete, more thorough, and more practical. People involved in developing and validating tests are better guided, and the result is more valid and better quality exams.

## References

- Bachman, L and Palmer, A (1996) *Language Testing in Practice*, Oxford: Oxford University Press.
- Cambridge English Language Assessment (2013) *Principles of Good Practice: Quality Management and Validation in Language Assessment* available online: [www.cambridgeenglish.org/research-and-validation/quality-and-accountability](http://www.cambridgeenglish.org/research-and-validation/quality-and-accountability)
- Field, J (2011) Cognitive validity, in Taylor, L (Ed), *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, Studies in Language Testing volume 30, Cambridge: UCLES/Cambridge University Press, 65-111.
- Field, J (forthcoming 2013) Cognitive validity, in Geranpayeh, A and Taylor, L (Eds), *Examining Listening: Research and Practice in Assessing Second Language Listening*, Studies in Language Testing volume 35, Cambridge: UCLES/Cambridge University Press, 77-151.
- Geranpayeh, A and Taylor, L (Eds) (forthcoming 2013) *Examining Listening: Research and Practice in Assessing Second Language Listening*, Studies in Language Testing volume 35, Cambridge: UCLES/Cambridge University Press.
- Hawkey, R (2009) *Examining FCE and CAE: Key Issues and Recurring Themes in Developing the First Certificate in English and Certificate in Advanced English Exams*, Studies in Language Testing volume 28, Cambridge: UCLES/Cambridge University Press.
- Kane, M (2006) Validation, in Brennan, R L (Ed), *Educational Measurement (4th Edition)*, Washington DC: American Council on Education/Praeger, 17-64.
- Khalifa, H and Weir, C J (2009) *Examining Reading: Research and Practice in Assessing Second Language Reading*, Studies in Language Testing volume 29, Cambridge: UCLES/Cambridge University Press.
- Lim, G S (2010) *Task options for a revised CPE Writing paper*, Cambridge ESOL internal report.
- Messick, S (1989) Validity, in Linn R L (Ed) *Educational Measurement (3rd Edition)*, New York: Macmillan, 13-103.
- Saville, N (2003) The process of test development and revision within UCLES EFL, in Weir, C J and Milanovic, M (Eds) *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913-2002*, Studies in Language Testing volume 15, Cambridge: UCLES/Cambridge University Press, 57-120.
- Saville, N (2012a) Applying a model for investigating the impact of language assessment within educational contexts: The Cambridge ESOL approach, *Research Notes* 50, 4-8.
- Saville, N (2012b) Quality management in test production and administration, in Fulcher, G and Davidson, F (Eds), *The Routledge Handbook of Language Testing*, Abingdon: Routledge, 395-412.
- Shaw, S and Weir, C J (2007) *Examining Writing: Research and Practice in Assessing Second Language Writing*, Studies in Language Testing volume 26, Cambridge: UCLES/Cambridge University Press.
- Somers, A, Geranpayeh, A and Malarkey, J (2010) *CPE construct investigation*, Cambridge ESOL internal report.
- Spillett, H (2012) The revision of the Cambridge English: Proficiency Writing paper, *Research Notes* 49, 2-5.
- Taylor, L (Ed) (2011) *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, Studies in Language Testing volume 30, Cambridge: UCLES/Cambridge University Press.
- Weir, C J (2005) *Language Testing and Validation: An Evidence-Based Approach*, Basingstoke: Palgrave Macmillan.
- Weir, C J and Milanovic, M (Eds) (2003) *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913-2002*, Studies in Language Testing volume 15, Cambridge: UCLES/Cambridge University Press.

# A look at test taker characteristics for a C2 exam

ARDESHIR GERANPAYEH RESEARCH AND VALIDATION GROUP, CAMBRIDGE ENGLISH LANGUAGE ASSESSMENT

## Introduction

Weir (in this issue) reports that only three candidates took the Certificate of Proficiency in English (CPE) in 1913. The test was targeted at measuring language proficiency of foreign students who wished to teach English in a foreign school. The test was later modified, as reported in Weir's article, to be relevant not only to candidates who wished to become teachers of English but also to those foreign students who would use English within a variety of academic disciplines. The target test population remains the same to this date, i.e., targeting mature students at college/university level. In this article, the changes to the CPE (now known as *Cambridge English: Proficiency*) population over the last two decades are reported as test taker characteristics were not consistently captured before the 1990s.

The importance of knowing test taker characteristics for test construction and administration has been well documented in the language testing literature (Bachman 1990, Geranpayeh and Taylor forthcoming 2013, Khalifa and Weir 2009, O'Sullivan 2000, Taylor 2011, Weir 2005, to name but a few). Who will take the test is a primary question that every test developer has in mind when constructing tests. Various physical/physiological, psychological and experiential characteristics of the test taker can play a role in how a candidate may respond to an item. Monitoring demographic changes to a test taker population allows examination boards to monitor unintended consequences of their test items for subgroups of test takers. This is often detected by employing Differential Item Functioning (DIF) methods in test analysis (see Geranpayeh 2008:16–23).

Weir (in this issue) lists the creation of the EFL Evaluation Unit (now the Research and Validation Group) in 1989 as one of the key events affecting the historical development of English language teaching in the United Kingdom and English language testing at Cambridge English Language Assessment. Prior to the creation of this unit, the information about CPE test taker characteristics was not consistently documented. We know that the test population grew significantly post-World War Two (and in particular after 1975) (Hawkey and Milanovic forthcoming 2013) but there is no documented literature to give us a consistent picture of the changes in the population prior to the 1990s. We assume that teachers and university students must have been 18 years and over but no other specific information is available. This deficiency was recognised by the EFL Evaluation Unit very early on in the 1990s and measures were put in place to collect such information for all Cambridge English exams, including CPE. A Candidate Information Sheet (CIS) was introduced in 1991 to capture test taker characteristics at the time of taking an examination. Cambridge English Language Assessment now holds millions of records of test taker characteristics over a period of 22 years. CIS data contains candidate information on various features such as their age, native language, nationality, education background, preparation courses and reasons for taking the test.

## CPE population 1991–2012

In order to investigate changes in candidate demography since the 1990s, we chose a summer session in June 1991 and compared that candidate CIS data with the latest summer 2012 candidate data. The June 1991 exam comprised around 25,000 candidates, 68% of which were female. Almost 86% of the population indicated that they had participated in test preparation courses prior to sitting the test. This is very much in line with other Cambridge English exams where they were integrated into school examination systems.

Figure 1 demonstrates that almost 46% of the candidature came from Greece. This is due to the Greek education system, which recognises CPE as a certificate of English proficiency for teaching English at schools. Other major L1 populations include French, Spanish, German and Italian. This picture remained very much the same until late into the 20th century.

Figure 1: CPE native language distribution in June 1991 (percentage)

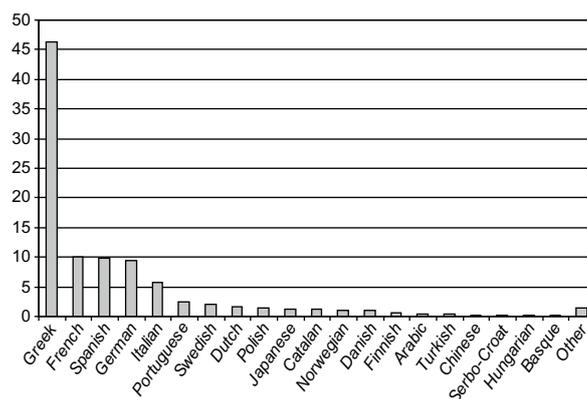
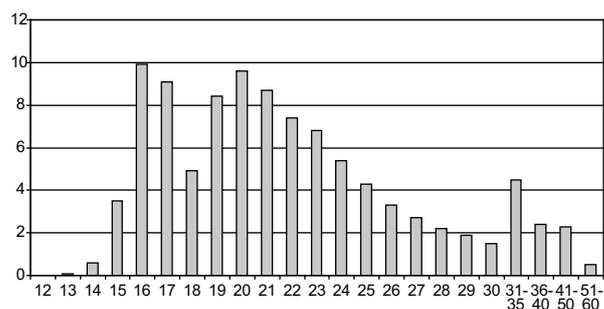


Figure 2 illustrates candidate distribution by age. The largest single age group is 16–20-year-olds who comprise 42% of the population, followed by 21–25 at 33% and 26–30 at 12%. The younger age group in this data is heavily influenced by the Greek candidates who take the exam at the age of 16. In fact, if we removed the Greek population from this data the proportion would be more like 34%, 34% and 17% for 16–20, 21–25 and 26–30 age groups, respectively.

Figure 2: CPE age distribution in June 1991 (percentage)



The information in Table 1 provides evidence that 30% of the population in 1991 took the test for employment purposes in

their own country. Only 10% showed interest in using it for employment abroad. It also demonstrates that the test results were mainly used in higher education and employment (e.g. teaching English) opportunities. What is unique about CPE 1991 data is the large number of candidates (27%) who took the test for personal interest. CPE was designed to show evidence of the highest level of English proficiency one can achieve as a foreign/second language, which is Level C2 according to the Common European Framework of Reference for languages (CEFR) (Council of Europe 2001). Many candidates took the test because they saw the CPE certificate as an ultimate record of their achievement in English language proficiency.

**Table 1: Reason for taking CPE in 1991 (percentage)**

Reason for taking CPE	Population*
Employment in your own country	30
Personal interest	27
Further study of English	21
Further study of other subjects	13
Employment abroad	10

\* percentages do not sum to 100 due to rounding.

CIS data from 1991–99 shows a very consistent picture of over several hundred thousand candidates who took the CPE test during that time. Information from the CIS data after 2000 influenced some of the decisions in test design which led to the 2002 CPE revision.

The population demographic starts to gradually change from 2007–08, moving away from a dominant Greek subgroup to a broader L1 population, but the candidature still remains very much European.

**Figure 3: CPE native language distribution in 2012 (percentage)**

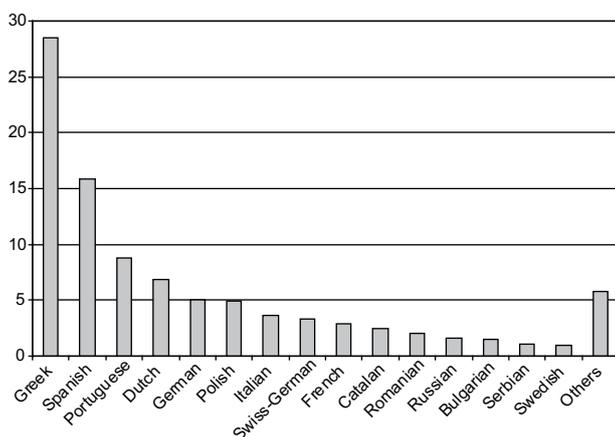
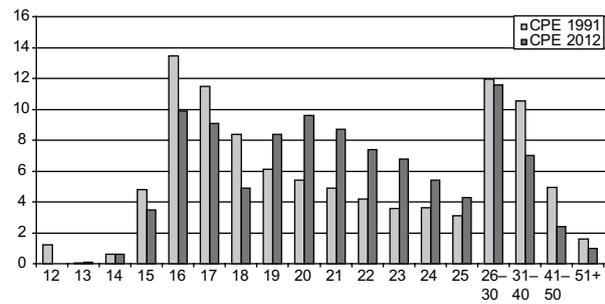


Figure 3 shows a breakdown of the CPE population in terms of native language reported. Although the Greek population is still the largest cohort, its dominance has significantly been reduced to just under 30%. The Spanish, Portuguese, Dutch and German cohorts have increased their share of the population, which shows CPE’s popularity across Europe. This should also have an impact on age distribution.

Figure 4 compares the CPE population of 1991 and 2012. Against common belief, it appears that the CPE population has not become younger; to the contrary, it has become more mature. The reason for this shift in population age is the

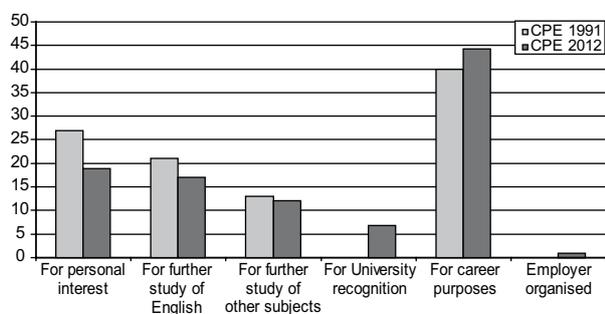
**Figure 4: CPE age distribution, 1991 vs 2012 (percentage)**



relative proportion of Greek candidates in 2012. The statistics for Greece and the rest of the world have not actually changed over the past 21 years. The Greek population is still very young compared to the rest of the world. What has changed is the popularity of CPE among non-Greeks, who tend to be more mature candidates, hence the proportion of mature candidates has changed in recent years. If we look at a country by country comparison, the CPE age group is still very similar to what it was in 1991.

Another change in CIS data since 1991 is the gender proportion. There is now a more balanced proportion between female and male candidates: 59% to 41%, respectively in 2012 compared to 68% to 32% respectively in 1991. It appears also that recently comparatively fewer candidates say they have attended preparation courses than in the past. However, 74% of candidates still claimed to have taken CPE preparation courses prior to taking the test in 2012. The 2012 CIS data shows that 55% of the population have a college or university education compared to 45% coming directly from secondary education.

**Figure 5: Candidate's reason for taking CPE, 1991 vs 2012**



In recent years, Cambridge has collected more detailed CIS data regarding the candidates’ reasons for taking the test. Figure 5 shows that a significant number of candidates still take CPE to help their career. Because of updates to the CIS forms, new information has been captured, and we can now see that a number of candidates are taking the exam to satisfy university entrance requirements. This category was not captured in the 1990s mainly because CPE certificates were judged to be well above university admission threshold levels, despite being recognised by most universities. However, the concept that CPE is considered to be a level too high for entrance to higher education is gradually changing, particularly for entrance to more linguistically challenging courses.

Another interesting statistic about the CPE population in 2012, is that candidates have indicated that on average they have been learning English for over six years prior to taking

CPE. Almost 60% of the population either claimed to have sat the CPE exam or other Cambridge English exams prior to taking the test. All of this shows that getting a CPE certificate is the final achievement of one's proficiency in English.

## Summary

The CPE (now known as *Cambridge English: Proficiency*) examination has always been a test of academic English since its inception in 1913. It targeted teachers of English and those who needed to demonstrate that they were at the highest proficiency level in English. Despite changes to its format over many years, its population remains very much at the top of the Cambridge English proficiency qualifications ladder. Cambridge CIS data shows that the test became very popular for many students in Europe towards the end of the secondary school curriculum during 1990–2000. The test population still remains very much European and linked to the English language teaching profession. It is a test well above the entry threshold required for many universities although it has always been academic in orientation.

## References

- Bachman, L F (1990) *Fundamental Considerations in Language Testing*, Oxford: Oxford University Press.
- Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.
- Geranpayeh, A (2008) Using DIF to explore item difficulty in CAE listening, *Research Notes* 32, 16–23.
- Geranpayeh, A and Taylor, L (Eds) (forthcoming 2013) *Examining Listening: Research and Practice in Assessing Second Language Listening*, Studies in Language Testing volume 35, Cambridge: UCLES/Cambridge University Press.
- Hawkey, R and Milanovic, M (forthcoming 2013) *Cambridge English Exams - the First Hundred Years: A History of English Language Assessment from the University of Cambridge, 1913–2013*, Studies in Testing volume 38, Cambridge: UCLES/Cambridge University Press.
- Khalifa, H and Weir, C J (2009) *Examining Reading: Research and Practice in Assessing Second Language Reading*, Studies in Language Testing volume 29, Cambridge: UCLES/Cambridge University Press.
- O'Sullivan, B (2000) *Towards a Model of Performance in Oral Language Tests*, unpublished PhD thesis, University of Reading.
- Taylor, L (Ed.) (2011) *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, Studies in Language Testing volume 30, Cambridge: UCLES/Cambridge University Press.
- Weir, C J (2005) *Language Testing and Validation: An Evidence-Based Approach*, Basingstoke: Palgrave Macmillan.

# Consulting stakeholders as part of the Cambridge English: Proficiency exam revision

**COREEN DOCHERTY** RESEARCH AND VALIDATION GROUP, CAMBRIDGE ENGLISH LANGUAGE ASSESSMENT

**DEBBIE HOWDEN** BUSINESS DEVELOPMENT GROUP, CAMBRIDGE ENGLISH LANGUAGE ASSESSMENT

Cambridge English Language Assessment regularly monitors the needs of test takers and the context of use of its exams to ensure its assessment products are fit for purpose and have a positive impact on learning (Cambridge English Language Assessment 2013). In 2009, Cambridge English Language Assessment, as part of its regular test development and revision process, conducted a review of the *Cambridge English: Proficiency* exam (see Zeronis and Elliott's article in this issue). A first and key step in the exam review process is to consult stakeholders to come to a better understanding of their views and ensure that the exam is meeting the needs of those that use the exam or make use of the results of the exam. Eliciting feedback from stakeholders is an ongoing process and this article aims to describe selected findings from two surveys conducted to investigate teachers' and centres' attitudes towards the *Cambridge English: Proficiency* exam both pre and post revision. The first survey was conducted at the beginning of the revision process in 2009 and the information gathered was used to inform test developers when making decisions concerning the revision of *Cambridge English: Proficiency*. The second survey was undertaken more recently to gather general perceptions of the new exam from centre administrators who were involved in pretesting it. It should be pointed out that these surveys represent only two examples of consultative exercises undertaken during the *Cambridge English: Proficiency* revision process, as stakeholder feedback was collected throughout the trialling stage to ensure that changes to the exam were fit for purpose (Weir and Milanovic 2003, see also Zeronis and Elliott's article in this issue). Selected responses to the pre-revision consultation will be presented below followed by the post-revision findings.

## Pre-revision consultation

In November 2009, 700 teachers and 170 centres in key countries where *Cambridge English: Proficiency* is taken were asked to submit their views about the current exam and factors that should be considered when modifying the exam. Responses were received from 153 teachers and 113 centres across 13 countries. The majority of the teachers (67.3%) who responded reported that they were currently preparing students to take the *Cambridge English: Proficiency* exam while 20.9% were not currently preparing students but had done so in the past.

### Factors influencing students to take Cambridge English: Proficiency

Teachers and centres were presented with 14 factors that could be seen to influence a learner's decision to take the *Cambridge English: Proficiency* exam and asked to indicate how

influential they felt each factor was. The factors that were deemed as the most influential are found in Table 1.

**Table 1: Main factors influencing students to take Cambridge English: Proficiency (percentage agreement)**

	Very and moderately likely	
	Teachers	Centres
It improves students' job prospects	94.8	92.9
It is recommended by the teacher/school	85.2	67.3
It is recognised for admission to universities in English-speaking countries	82.2	78.7
It is a prestigious exam	80.8	77.8
It is for students' personal development	75.5	73.5
To demonstrate students have achieved a higher level of English than CAE	74.8	77
It is recognised in some countries for teaching purposes	72.6	64.6

Although the factors considered as most influential are the same for both groups of respondents, teachers thought that teacher/school recommendations and the need for *Cambridge English: Proficiency* for teaching purposes were more important than did centres. This may be a result of centre administrators being less aware of classroom factors. In response to an open-ended comment option, five teachers suggested that students are also influenced to take the exam if they had previously taken other Cambridge English exams, highlighting candidates' desire to climb to the top of the Cambridge English qualifications ladder. Examples of typical responses are as follows:

From a teacher in Greece:

Some students and parents believe that this is a natural progression after FCE [*Cambridge English: First*] and CAE [*Cambridge English: Advanced*] and will lead to a completion of their English studies.

From a teacher in Bulgaria:

Due to the fact that students have taken FCE, CAE and they want to have a set of the three most prestigious certificates.

Centre comments primarily focused on students taking the *Cambridge English: Proficiency* exam for work-related purposes; for example:

From a centre in Spain:

A very high level of fluency is required in certain jobs, for example, translation, academic or business. Candidates are usually working in professions where they are dealing with native speakers and need to understand everything and be able to write extremely well.

A few centres also noted that the increased presence of English in schools and higher education has resulted in more learners at the level of *Cambridge English: Proficiency*:

From a centre in Spain:

[With the] higher level of language skills and increased study time of younger learners, we are beginning to see the effects of starting English at a much earlier age in Spain. We are also seeing the effects of Erasmus programmes, studying abroad, etc. CPE [*Cambridge English: Proficiency*] is seen as the “final” most advanced certificate – as 10 years ago few people had FCE, now many students [and] teachers [have] FCE (even CAE). CPE is seen as “superior” to these two certificates.

Finally, some centres pointed out that students take the exam solely for personal satisfaction:

From a centre in Italy:

[They take the exam for] the sheer love of the language, instilled largely by the driving force of teachers.

### Factors to be considered when updating Cambridge English: Proficiency

Teachers and centres were then asked which factors should be considered when updating the *Cambridge English: Proficiency* exam. A summary of the responses is found in Table 2.

**Table 2: Factors to be considered when updating Cambridge English: Proficiency (percentage agreement)**

	Very important and important	
	Teachers	Centres
Content suitable for general purposes	79.7	62.8
Content suitable for study/academic purposes	78.5	87.6
Content suitable for a younger age group (under 18 years old)	77.7	54.9
Content suitable for career advancement	75.2	87.6
Students who achieve a ‘narrow fail’ at <i>Cambridge English: Proficiency</i> should be awarded a CEFR Level C1 certificate	73.2	69.9
A computer-based version	61.4	53.1
<i>Cambridge English: Proficiency</i> exam should become shorter	54.5	61.9

Teachers placed more emphasis on the content being suitable for both general purposes and a younger age group than did centres. However, as Geranpayeh (this issue) points out, candidate information suggests that *Cambridge English: Proficiency* test takers are not getting younger. This contradiction between stakeholder perceptions and test taker data highlights the importance of using a variety of data sources during a revision process.

Both teachers and centres agreed that any revision to the exam should include content suitable for academic purposes and career advancement. Teachers were given the opportunity to describe the features of an effective test of English for study purposes and about one third of the 102 responses indicated that this type of test should contain academic-style tasks, such as note-taking, listening to lectures, writing essays and summarising information. An example of a typical comment is as follows:

Teacher in Brazil:

For study purposes an effective test should have a focus on academic skills, where reading and writing play an important role. The testing of vocabulary and structures related to the academic world are also essential, since these are mandatory for good comprehension and writing skills. As for the speaking exam, I believe the exam must test skills that will allow the student to take an active part in the academic institution. The listening

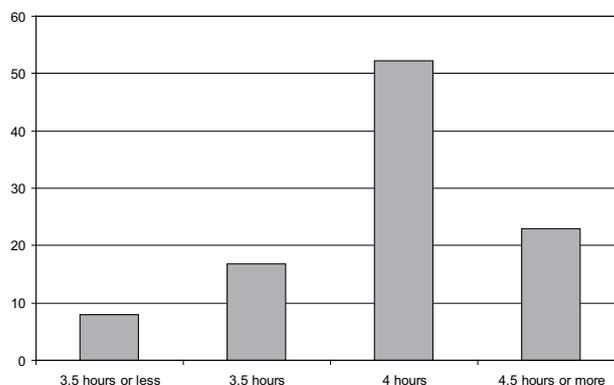
should focus on the understanding of lectures and talks since that’s what students will be more exposed to.

Other factors considered important when updating the examination included offering candidates who narrowly fail the exam a certificate at the C1 level of the Common European Framework of Reference for languages (CEFR) and offering a computer-based version of the exam.

### Optimal timing of a C2-level exam

The appropriate length of time needed to demonstrate competence in a CEFR Level C2 exam was investigated. The combined time needed to complete all five papers in the current version of the *Cambridge English: Proficiency* exam is just under 6 hours and there was a concern that this length of time may be leading to candidate fatigue. Figure 1 shows that a 4-hour exam was considered the most appropriate length for a C2 exam followed by 4.5 hours or more and then 3.5 hours.

**Figure 1: Optimum length of time to demonstrate C2-level competence (percentage agreement)**



Although there was a clear preference for shortening the time taken to complete the exam, comments suggest that any change in timing should not compromise the quality of the exam or be seen to lower the level of difficulty. An example of a typical response is as follows:

From a teacher in Brazil:

I believe the CPE exam could be shortened, but not much . . . and should remain a high-level and complex exam.

### Pre-revision conclusions

The pre-revision consultative exercises highlighted a number of key areas that needed to be considered during the revision process. In particular, that teachers and centres believe the primary reasons test takers take the *Cambridge English: Proficiency* exam are to improve their job prospects, to gain admission to university or because they believe it is a prestigious exam. These factors were also considered important when respondents were asked which factors should be considered when updating the exam. They pointed out that the content should be suitable for both study purposes and career development but also that the qualifications should recognise candidate performance that is just below the C2 level by awarding those candidates a CEFR C1 certificate. A test of approximately 4 hours was considered an appropriate length for candidates to demonstrate C2-level language competence. A shorter test, however, should not come at the expense of reducing the overall quality or level of difficulty of

the qualification which stakeholders value and is linked to its recognition as a prestigious exam. This point is highlighted by a teacher in Greece:

I believe that CPE is for those few who do not seek superficial knowledge and so will spend time preparing. CPE is a certificate that pushes students into realising their full potential, makes high demands on them and doesn't adapt its level to the students but aims to bring them to the level that a CPE holder should have. The difference is that whoever holds CPE knows English, which is not always the case with the other certificates.

The views of stakeholders were taken into consideration during the revision process resulting in an exam with an enhanced academic focus, a shorter timing, a computer-based version and enhanced certification<sup>1</sup>, among others (see Zeronis and Elliott's article in this issue for a full description of changes made to the *Cambridge English: Proficiency* exam).

## Post-revision consultation

In September 2012, 268 Centre Exams Managers who had pretested<sup>2</sup> the new *Cambridge English: Proficiency* exam in the previous six months were sent a questionnaire to investigate their general attitude towards the changes made to the exam. Eighty-eight responses were received from centres with the majority of respondents indicating they were teachers (54.5%).

In general, the feedback received from respondents on the new exam was encouraging with 93% of them indicating that they had a positive or very positive perception of the new *Cambridge English: Proficiency* exam. As many respondents who participated in the pre-revision consultative exercises indicated the importance of ensuring the exam was suitable for study purposes, career advancement and candidates under the age of 18 (see Table 2), the post-revision respondents were asked if these aims had been achieved.

**Table 3: Appropriacy of the new *Cambridge English: Proficiency* exam for different stakeholders (percentage agreement)**

	Strongly agree/agree	Strongly disagree/disagree	Don't know
All ages 16+	57.8	38.6	3.6
Undergraduates	76.6	23.4	0
Postgraduates	98.7	0	1.3
Professionals*	94.7	2.7	2.7

\* percentages do not sum to 100 due to rounding.

Table 3 shows that the exam is perceived to be most appropriate for postgraduates and professionals followed by undergraduates and lastly by all ages 16 and over. It is not surprising that the least amount of agreement is related to the appropriateness of the exam for all ages. A key feature of the C2 level is that learners are able to handle abstract ideas and concepts, which requires a certain level of cognitive

development or maturity that a younger candidate (under the age of 18) may not possess (Council of Europe 2001, Ramshaw 2010). The following comments from respondents highlight this issue:

From a teacher in Uruguay:

In our country students are not ready to take CPE before 18, in general and sometimes not even then due to cultural constraints or lack of intellectual maturity.

From a teacher in Brazil:

I am looking forward to the new changes - I like the idea of making it shorter in terms of time. But, it is a test that is not for very young people - I had two 15-year-old candidates for Proficiency last year: they know a lot of English but were not so mature for the compositions - thus, it can be a problem.

Although the academic focus of the exam was enhanced (see Zeronis and Elliott's article in this issue), respondents did not feel its accessibility was limited as there was 90% agreement that the exam is appropriate for all nationalities and cultures.

When asked to compare the new *Cambridge English: Proficiency* exam with the current exam, the response was positive with 80.2% of respondents indicating that the new exam is either equal to, better or much better than the current exam (see Table 4).

**Table 4: Comparison between the new format of the *Cambridge English: Proficiency* exam and the current one (percentage agreement)**

	Much better/better	Equal	Worse/much worse	Don't know
Compared to the current exam, the new exam is . . .	68.6	11.6	7.0	12.8

Respondents were given the opportunity to explain why they thought the exam was better or much better, and the most popular response was the shorter length of time needed to complete the exam in comparison to the current exam. For example:

From a teacher in the Ukraine:

It saves much time, it is physically easier for candidates to withstand it in only 4 hours, and the quality of evaluation is still high.

From a teacher in Uruguay:

In a shorter time the candidates can demonstrate their skills in the same way as in the previous format, but more effectively and economically. This makes the exam more user-friendly.

Other respondents commented on improvements to task types:

From a teacher in the Former Yugoslav Republic of Macedonia:

First of all because of the new type of listening and writing tasks - according to me as a teacher, those parts are for the students who can prove that their English is on a rather high level - saying that those students can live, work, study in a country where English is an official

<sup>1</sup> Cambridge English Language Assessment is now offering enhanced certification for certain exams, which means that if a candidate who, for example, is taking a B1-level test performs well enough to demonstrate B2 ability, he/she will be awarded a B2 certificate. Similarly if a candidate taking a B1 test does not perform at the B1 level but does demonstrate A2 ability, he/she will be awarded an A2 certificate.

<sup>2</sup> Pretesting is a stage in the question paper production process where tasks are trialled on learners who are representative of the population who normally take the live test. Pretesting allows Cambridge English Language Assessment to determine the effectiveness of the test material and establish measurement characteristics but it also gives candidates an opportunity to practise taking the test they are preparing for under exam conditions and receive feedback on their performance.

language. Second thing that the students like better is the fact that the whole exam is 4 hours in total.

From a teacher in Spain:

Because it is much more concise than anything and its similarity to CAE encourages students to continue their studies.

Of those that felt the new exam was equivalent to the current exam, the majority of respondents mentioned that the overall level of difficulty of the exam had been maintained despite the change in tasks and reduction in timing:

From an examiner in the Ukraine:

Although some of the tasks have been changed, the complexity of the test seems to be very similar. I regularly take CPE pretests myself and I cannot say that the new version[s] (or my scores) are too different. However, I didn't get as exhausted as I used to get with the older version[s].

The respondents who felt the new exam was worse or much worse than the current exam made comments related to the loss of a favourite task, lack of familiarity with a new task or lack of available preparation material. These comments highlight the importance of communicating changes to stakeholders. Communication helps to ensure that tests have a positive impact on users and as such is an aspect of the Cambridge model of 'impact by design' (Cambridge English Language Assessment 2013:31-32). In the months leading up to the launch of the new *Cambridge English: Proficiency* exam, a number of events have been planned or have already taken place to increase stakeholder awareness of the changes to the exam. It is hoped that these awareness-raising campaigns will provide the information stakeholders need as well as give them an opportunity to provide more feedback.

## Conclusion

Consultative exercises are an essential part of not only the revision process but also the routine test production process as they allow test developers to monitor their products to ensure they are fit for purpose. The pre-revision findings described here provided useful information on the characteristics of *Cambridge English: Proficiency* test takers and the contexts of use which, when combined with information collected from test takers themselves and other stakeholders such as receiving institutions, helped guide test developers when making changes to the exam. The post-revision findings allowed test developers to assess the response to the changes made to the exam and address any issues that were raised. It is hoped that this article helps in demonstrating the importance of investigating stakeholders' views and how test developers can use this information to guide revision.

## References

- Cambridge English Language Assessment (2013) *Principles of Good Practice: Quality Management and Validation in Language Assessment* available online: [www.cambridgeenglish.org/research-and-validation/quality-and-accountability](http://www.cambridgeenglish.org/research-and-validation/quality-and-accountability)
- Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.
- Ramshaw, G (2010) *Expert Group Report for Award Seeking Admission to the UCAS Tariff: University of Cambridge ESOL Examinations, Certificate of Proficiency in English (CPE) and Certificate in Advanced English (CAE) grade A*, available online: [www.ucas.ac.uk/documents/tariff/tariff\\_reports/cambridgeesolreport.pdf](http://www.ucas.ac.uk/documents/tariff/tariff_reports/cambridgeesolreport.pdf)
- Weir, C and Milanovic, M (Eds) (2003) *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913-2002*, Studies in Language Testing volume 15, Cambridge: UCLES/Cambridge University Press.

# Development and construct of revised Cambridge English: Proficiency

RON ZERONIS ASSESSMENT AND OPERATIONS GROUP, CAMBRIDGE ENGLISH LANGUAGE ASSESSMENT

MARK ELLIOTT RESEARCH AND VALIDATION GROUP, CAMBRIDGE ENGLISH LANGUAGE ASSESSMENT

## Introduction

Cambridge English exams are reviewed at regular intervals and periodically updated to ensure that they are valid and useful instruments for test users. Thinking about the nature of language ability and how it should be taught and tested changes, and tests need to reflect this. At the same time, the candidates that take the test and the reasons they take these tests may also change, and tests need to account for these as well. A test development, review and revision process, as described in Saville (2003), is followed, and a result of this is an updated version *Cambridge English: Proficiency* from 2013.

The new *Cambridge English: Proficiency* exam represents an update rather than a major revision. As the history of Weir, Vidaković and Galaczi (forthcoming 2013) shows, the same dominant view of language learning and teaching has held sway since the 1970s, which sees language as a tool for communication. The Communicative Approach naturally invites focus on each of the four skills, and the four skills plus Use of English division had been adopted in the 1975 revision of the exam (Weir 2003). There being no major change in the underlying construct since then, subsequent revisions in 1984 and 2002 have thus been exercises in sharpening the focus of the test for the candidature it serves. The same is true for the revision leading to the 2013 version of *Cambridge English: Proficiency*.

The world has changed much since the 2002 modifications, the work for which actually began in 1992 (see Weir and Milanovic 2003). Communication now happens across distances and using media unimaginable then, which ultimately impacts on the way we communicate. For this reason, finding out users' needs is crucial, hence the importance of consultations with stakeholders. Part of the consultation process which informed this revision of *Cambridge English: Proficiency* is described by Docherty and Howden (this issue) and the demographics of recent *Cambridge English: Proficiency* candidates were also analysed (see Geranpayeh, also this issue). The analyses revealed, among other things, that candidates are taking the test for study purposes and there is a desire for a shorter test. The potential introduction of a computer-delivered version of *Cambridge English: Proficiency* also necessitated reviewing the test to ensure that task types are compatible with this mode of delivery.

In the European context, the changes were exemplified by unprecedented economic integration which highlighted the need for cross-cultural communication. It is in this context that the Council of Europe (2001) produced the Common European Framework of Reference for languages (CEFR), with its six reference levels of language ability from A1 to C2, and since then it has become an important document that language teachers and testers are increasingly relying on. While the C2 level of the CEFR was based on CPE (*Cambridge*

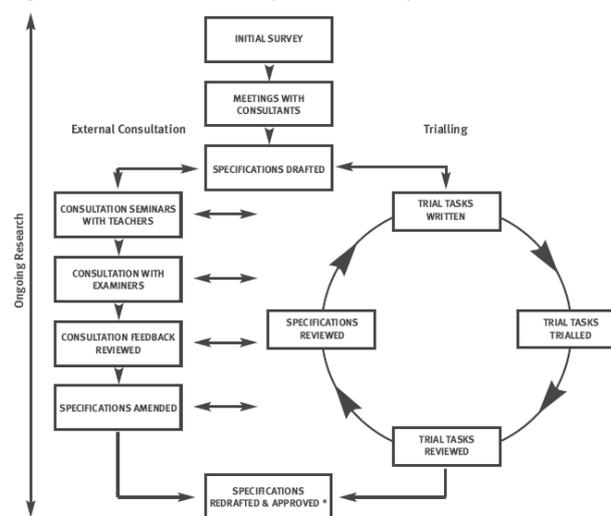
*English: Proficiency*) (North 2004), this revision allowed Cambridge English Language Assessment to review and make more explicit the relationship between the reference framework and the test.

This update thus provided Cambridge English Language Assessment with an opportunity to redevelop a test in line with current needs and realities while still maintaining a communicative construct. Specifics of those changes and what they address are detailed in the rest of this article.

## The revision process

Using the test development model described in Saville (2003) as illustrated in Figure 1, the project to review *Cambridge English: Proficiency* was initiated in 2009. This process began with a number of research and consultation exercises, including detailed reports on all aspects of the exam by senior examiners and materials developers, detailed performance reports on the current exam by the Research and Validation Group and consultative exercises involving centres, teachers and the organisation's regional business development staff. Throughout the process, we strove to maximise the appropriate balance of the following qualities: validity, reliability, impact, practicality and quality (VRIPQ) which underlie our approach to language assessment (see *Principles of Good Practice: Quality Management and Validation in Language Assessment* for more information on VRIPQ (Cambridge English Language Assessment 2013)).

Figure 1: Model of the revision process (from Spillett 2012:2)



Over 700 teachers and 170 centres around the world were surveyed as part of the consultation process (see Docherty and Howden's article in this issue for more information). Analysis showed that both groups felt that:

- content in the revised *Cambridge English: Proficiency* should be suitable for study purposes and career advancement
- the exam should be shorter (many felt that some individual papers contained too many tasks)
- a computer-based version would be popular
- the testing focuses and construct coverage in the current exam was about right.

In early 2010, based on the above research and feedback, a set of draft specifications and sample material was developed. Following a number of consultation meetings with stakeholders, amendments were made to the specifications and test versions for trialling were developed.

Several rounds of developmental trialling took place throughout 2010 and early 2011 in a number of countries all over the world – covering Europe, South America and Asia – on a significant number of candidates. Early rounds of trialling for the 2013 *Cambridge English: Proficiency* exam focused on the new task types being introduced into the Reading, Writing and Listening papers, as well as time trialling to ensure the length of each paper was appropriate. This was followed up by trials centred on task instructions and length of texts candidates produce in the new Part 1 task in the Writing paper (Spillett 2012). Additional trialling was also done on the order of tasks and numbers of items in the new Reading paper. Further time trialling was then carried out, with the final round completed at the end of January 2011. Throughout the process, the development team carried out in-depth analysis of the results, as well as held consultation meetings with stakeholders to ensure the new exam would meet the needs of users.

In April 2011, the final specifications were released along with a full set of sample papers for the revised exam. The commissioning of live material began and attention turned to internal systems development work needed to support the new exam and the development of the computer-based version of the test. In April of the following year, a full *Handbook for Teachers* (Cambridge ESOL 2012) was released, with a further set of sample papers, guidance for preparing candidates, and sample writing scripts.

## The new Cambridge English: Proficiency exam

From March 2013, *Cambridge English: Proficiency* will move from the current 5-paper format to a new 4-paper format where the Reading and Use of English papers are combined into one paper (see Appendix 1 to compare the previous format with the new format). The total timing of the paper has also been reduced from just under 6 hours to just under 4 hours by removing or modifying tasks which have similar testing focuses. A computer-delivered version will also be offered in 2013. A detailed description of the changes made to each paper follows.

### The Reading and Use of English paper

Paper 1 in the 2013 *Cambridge English: Proficiency* exam is a combination of the Reading and Use of English papers into a new, single paper (see Table 1). The paper retains a focus on reading skills (Parts 5–7), as well as tasks with a clear focus on grammar and vocabulary knowledge (Parts 1–4). Many of the tasks found in both the current Reading and Use of English papers have been retained in modified forms, and a new multiple-matching Reading task has been introduced (see Appendix 2 for more details on the structure of the paper and tasks).

The decision to combine Reading and Use of English was taken as these two papers share a number of key features. With the single exception of the key word transformations (Part 4), which are sentence-based, all tasks in the new paper are text-based. The Use of English texts (Parts 1–3) are of course somewhat less complex than the Reading texts (Parts 5–7) and naturally much shorter. However, there are broad similarities between Use of English texts and Reading texts as they both require reading comprehension and there are similarities in terms of the manner in which candidates read and process (see Khalifa and Weir 2009). For example, candidates are advised first to read through the Use of English texts quickly and to try to follow a line of argument – as with Reading tasks. In addition, previous research has traditionally shown that tests of language knowledge (vocabulary and grammar) correlate highly with tests of reading comprehension (see Alderson 2000, Read 2000).

Other changes to the paper include the addition of a multiple-matching task (Part 7). As this task also appears

**Table 1: Summary of the Reading and Use of English paper**

Paper/timing	Content
<b>Reading and Use of English</b> 1 hr 30 mins	<p><b>Part 1</b> A multiple-choice task consisting of a short text with eight gaps followed by eight multiple-choice questions. Candidates must choose one word or phrase from a set of four to fill each gap. The focus is on vocabulary.</p> <p><b>Part 2</b> An open cloze task consisting of a short text with eight gaps. Candidates think of the word which best fits each gap. The focus is on awareness and control of grammar and vocabulary.</p> <p><b>Part 3</b> A word formation task consisting of a text with eight gaps. Each gap corresponds to a word. The stems of the missing words are given beside the text and must be changed to form the missing word. The focus is on vocabulary, in particular the use of affixation, internal changes and compounding in word formation.</p> <p><b>Part 4</b> Key word transformations. The task consists of six discrete items with a lead-in sentence and a gapped response to complete in three to eight words including a given 'key' word. The focus is on grammar and vocabulary.</p> <p><b>Part 5</b> A Reading task consisting of a long text, followed by six 4-option multiple-choice questions. The task tests understanding of detail, opinion, attitude, tone, purpose, etc.</p> <p><b>Part 6</b> A Reading task consisting of a text from which paragraphs have been removed and placed in jumbled order after the text. Candidates must decide from where in the text the paragraphs have been removed. There are seven paragraphs, plus one extra which doesn't fit in the text.</p> <p><b>Part 7</b> A multiple-matching Reading task consisting of a single text of several short texts, preceded by 10 questions. Candidates must match each question to the correct text. The task tests understanding of detail, opinion, attitude, specific information, etc.</p>

on the *Cambridge English: First* and *Cambridge English: Advanced* exams, candidates are given a sense of continuity and progression across levels, which is something that was identified as positive by stakeholders (see Docherty and Howden's article in this issue). However, the ratio of questions to text length is slightly different from *Cambridge English: First* and *Cambridge English: Advanced* in that there are fewer questions over longer texts thereby better facilitating the skimming/scanning focus of the task (Khalifa and Weir 2009). Advantages of including this task are: this type of reading is widely used in both work-related and academic settings; this task type provides a wider variety of and greater distinction between tasks within the paper and has allowed the introduction of a set of reading focuses which were not previously included in the paper. It is possible for informal 'spoken' texts, such as four people being interviewed, to be represented in matching tasks. This type of text, when used, will lead to greater variety of text style, thereby compensating for the removal of the old short multiple-choice task from the new Reading and Use of English paper. The introduction of this task type also increases the task variety and response formats. It has been suggested that different response formats measure different aspects of language ability (Alderson 2000, Alderson, Clapham and Wall 1995, Khalifa and Weir 2009). Therefore, having a range of response formats could ensure that the test is not biased towards a particular kind of candidate (Alderson et al 1995) and that all candidates have the opportunity to perform at their best.

In order to avoid duplication of constructs across tasks and reduce the overall length of the paper, tasks were removed from the Reading paper and from the Use of English paper. The old Reading Part 2 task, short multiple-choice texts, happens to have the same task focus and response format (i.e. 4-option multiple choice) as Reading Part 4, the long multiple-choice texts (Part 5 in the 2013 version), with the focus being on the understanding of 'detail, opinion, attitude, tone, purpose, main idea, implication and text organisation features' (Cambridge ESOL 2008:7). A content analysis based on the reading comprehension model discussed in Khalifa and Weir (2009) reveals that the two test parts also elicit the same type of reading: careful reading across sentences and paragraphs in order to comprehend main ideas (2009). These test parts were also shown to activate similar cognitive processes, testing the ability of candidates to integrate new and old information (2009:73) and thus construct the meaning representation at the level of several sentences, paragraph(s) or the entire text. Although short and long multiple-choice texts can be very different in terms of style, source and purpose, the testing focus and task format of both these tasks is the same and for these reasons it was decided to remove this task type from the new paper.

Part 3 of the Use of English paper, gapped sentences, was also identified as a task that could be removed. The focus of Part 3 is on collocations, phrasal verbs and other word combinations. Each of the six questions consists of three discrete sentences and each sentence contains a gap. A candidate is required to fill the gaps with a single word common to all three sentences (Cambridge ESOL 2008:36). The task gives candidates the opportunity to show their awareness of how the senses and usage of a single word can vary with context. The task was excluded from the new

exam because of the overlap in task focus with the previous lexical cloze task in Part 1 Reading which, besides testing the meaning of individual words also tests 'fixed phrases, idioms and collocations' (Cambridge ESOL 2008:8). It is, therefore, believed that the Use of English gapped sentences task focus is 'adequately covered in the lexical cloze' (Cambridge ESOL 2011b:4). This is supported by the following two reasons: 1) Use of English Part 3 has the same kind of lexical task focus as the previous Reading Part 1 which has been retained in the new *Cambridge English: Proficiency* Reading and Use of English paper and 2) other Use of English test parts in the Reading and Use of English paper are expected to activate the same kinds of cognitive processes and types of reading as Use of English Part 3, in addition to the capacity to encourage reading above sentence level. In short, there seems to be nothing to suggest that excluding Use of English Part 3 from the Reading and Use of English paper would reduce construct coverage.

The testing focus of the comprehension questions from Use of English (the old Part 5) were also found to overlap with other tasks and the summary writing question from this task was thought to be better placed in another paper. The questions in this part focus on the gist of a written text and its overall function and meaning; inference; paradox; specific detail; the force of lexical items or phrases; and organisational features of a text, for example comparison and contrast, exemplification and anaphoric/cataphoric reference (Cambridge ESOL 2006). These testing focuses overlapped considerably with Part 2 of the old Reading paper (Part 5 of the new Reading paper), which included in its testing focus: text content features (detail, opinion, attitude, tone, writer's purpose, main idea and implication) and text organisation features (exemplification, comparison and reference) (Cambridge ESOL 2006:12). For this reason, the comprehension questions have been removed from the revised test to avoid the overlap without incurring any significant effect on construct coverage. The summary writing task has been expanded and moved to the Writing paper, as discussed below.

#### *Sequencing of tasks*

Finally, following extensive trialling, it was decided to put the shorter, text-based tasks first (the multiple-choice cloze (Part 1), open cloze (Part 2), and word formation (Part 3) are standardised at eight items and have a similar word length). This provides the same 'easing in' effect that the old Reading paper had; i.e. texts go from short to long, and increase in complexity. Trialling strongly suggested that if the Reading tasks were to appear first, weaker candidates, in particular, might well spend longer on what are long and complex texts, leaving themselves insufficient time to fully do justice to their ability in the Use of English section.

#### *Overall*

The merging of the Reading and Use of English papers has not essentially changed the construct. The current level and breadth of coverage has been maintained and it is still possible to clearly distinguish use of English tasks from Reading Comprehension tasks in the new Reading and Use of English paper, based on task focus, format and the levels of cognitive processes they are predominantly expected to activate. Some support for merging the old Reading and Use of English

papers into one paper was found in the fact that they are expected to activate cognitive processes along the continuum of the reading comprehension model discussed in Khalifa and Weir (2009). In addition, previous research has shown the close link between linguistic knowledge and reading comprehension: 'Readers rely more on linguistic skills, both decoding and linguistic comprehension, to decode words and to comprehend sentences accurately, so that salient points can be synthesised to build up main ideas and propositional and pragmatic inferences can be made' (Weir and Jin 1996). Similarly, Weir and Porter (1994:8) rightly argue 'it does seem improbable that students would be able to work out the main ideas of a text without some baseline competence in the microlinguistic skills'.

The new Reading and Use of English paper is more concise without sacrificing quality, difficulty or coverage. It is almost all text based and continues to assess candidates' abilities through the use of a range of different texts from a variety of sources, making it an even more suitable and comprehensive assessment tool for both higher education study and career advancement purposes.

### The Writing paper

The Writing paper also underwent changes including a new compulsory summary task (Part 1), which is a modified version of the Part 5 summary question from the old Use of English paper. A minor change was made to the writing options elicited in Part 2 as well. The overall timing of the Writing paper has also been reduced from 2 hours to 1 hour 30 minutes. Table 2 provides an overview of the new Writing paper and Appendix 3 provides more information on the structure of the paper.

**Table 2: Summary of the Writing paper**

Paper/timing	Content
<b>Writing</b> 1 hr 30 mins	<p><b>Part 1</b> One compulsory question. Candidates produce an essay with a discursive focus. The task requires candidates to summarise two short texts, and then to evaluate the ideas contained in them.</p> <p><b>Part 2</b> One question from a choice of five (including the set text options). Candidates produce a text in one of the following genres: article, letter, report, essay, or review.</p>

The new compulsory Part 1 question requires candidates to read two short passages of approximately 100 words each on a particular topic. The texts may contain complementary or contrasting opinions, and may be extracts from newspapers, books, magazines, online source material, or could be based on quotations made by speakers during a discussion. Candidates are then to respond to the input material by writing an essay which summarises the key points from both texts and evaluates the same. This task is highly appropriate for a test of writing at the C2 level as it is in line with current thinking on the importance of being able to deal with intertextual meaning at the highest levels of language ability (Shaw and Weir 2007, Weir and Milanovic 2003). At this level, writing is agreed to be a complex, problem-solving activity involving significant cognitive resources. This is captured, for example, in Bereiter and Scardamalia's (1987) conceptualisation of writing that requires knowledge transformation.

In addition, in academic and other contexts, the material to be transformed often involves source materials (e.g. course readings, official documents), and writing is about relating those texts with each other and with one's own ideas (Carson and Leki 1993). The same ideas are reflected in the CEFR's illustrative descriptors for the C2 level. They state that candidates at this level should be able to 'write complex letters, reports or articles which present a case with an effective, logical structure which helps the recipient to notice and remember significant points', and produce '... summaries and reviews of professional or literary works' (Council of Europe 2001:27). Elsewhere in the CEFR, it adds that candidates should be able to 'summarise information from different sources, reconstructing arguments and accounts in a coherent presentation of the overall result' (2001:96). In view of the above, a writing task that requires summarising ideas from multiple texts and engaging with the ideas contained in them provides a way of engaging the cognitive processes involved and eliciting writing samples that reflect the outcomes of that complex activity.

Indications from trialling of the new task have been positive. Trialling scripts have been marked in conjunction with the new analytical mark scheme aligned to CEFR level descriptors for writing proficiency. Evidence shows that candidates at C2 level are able to identify key points and develop these in conjunction with their own ideas in an essay with a discursive focus. Moreover, the indicative word range of 240–280 words provided sufficient scope for C2-level candidates to respond well to the new Part 1 task. However, the strongest candidates consistently produced responses of more than 300 words. It should be noted that candidates are not penalised for exceeding the recommended word lengths (unless indirectly if additional text leads, for example, to irrelevance or incoherence), and should therefore feel free to provide longer responses if that allows them to respond to the task best.

Although there has been a major change to Part 1 of the Writing paper, Part 2 remains largely unchanged. This part no longer includes the option of writing a proposal, but has retained the report option (which involves functions similar to those in writing a proposal), as well as the options of an article, a letter, a review and, in the case of the two set text options, the additional option of an essay task. The retention of set texts addresses the provision of contemporary literature as a means of engaging with language at C2 level. Many students and teachers alike testify to the benefits gained by studying a set text at this level, even if that option is not taken up in the actual examination (Fried-Booth 2004).

After four rounds of trialling (with candidates across the globe), it was concluded that a guideline range of 280–320 words for Part 2 questions provided appropriate scope for C2-level candidates to respond to the whole Writing paper in the reduced allotted time. As with Part 1, candidates exceeding the suggested upper limit are not penalised under the aligned mark scheme (see Lim 2012 for more information concerning the aligned mark scheme).

### Overall

The new Writing paper ensures a consistency of approach in maintaining comprehensive coverage of writing skills at C2 level, while simultaneously reducing the timing of the paper and length of the required output. Although the reading input

**Table 3: Summary of the Listening paper**

Paper/timing	Content
<b>Listening</b> 40 mins (approx.)	<p><b>Part 1</b> Candidates listen to three short, unrelated texts lasting approximately 1 minute each, consisting of either monologues or exchanges between interacting speakers. There are two 3-option multiple-choice questions per extract. The task focus is on identifying speaker feeling, attitude, opinion, etc.</p> <p><b>Part 2</b> Candidates listen to a monologue (which may be introduced by a presenter) lasting 3 to 4 minutes. Candidates are required to complete nine sentences with information heard on the recording. The task focus is on identifying specific information and stated opinion.</p> <p><b>Part 3</b> Candidates listen to an interview, or a conversation between two or more speakers of approximately 4 minutes. There are five 4-option multiple-choice questions. The task focus is on identifying attitude and opinion, gist, detail, etc.</p> <p><b>Part 4</b> Candidates listen to five short, themed monologues of approximately 35 seconds each. There are two tasks. Each task contains five questions and requires selection of the correct option from a list of eight. The task focus is on identifying gist, attitude, main points, and interpreting context.</p>

for Part 1 has been increased, candidates are required to write fewer words. Furthermore, the shift in Part 1, which is designed to synthesise reading, summary and writing skills, provides the opportunity to prepare more effectively for university-level study and career development (for a more detailed analysis of the new *Cambridge English: Proficiency Writing* paper, see Spillett 2012).

### The Listening paper

Changes to the Listening paper include one fewer short extract task in Part 1, and the matching task in Part 4 (3-option multiple matching based on one informal discussion) has been replaced with a multiple-matching task based on five short monologues similar to that of *Cambridge English: Advanced* (see Table 3 for a summary of the paper and Appendix 4 for more detailed information on the structure of the paper and tasks).

Although the format of Part 1 has not changed, the number of tasks has been reduced from four to three and the number of questions from eight to six. As the testing focuses of Parts 1 and 3 are broadly similar, with the main difference being text length and genre, it was decided that Part 1 could be reduced without significantly affecting the range of what is being tested whilst maintaining a good variety of interaction patterns across the two tasks. Despite the reduction in the number of items in Part 1, the overall reliability of the test has been maintained with the introduction of the new Part 4 task, which increases the total number of items in the test from 28 to 30.

Part 2 remains a monologue of an informative nature, with a sentence-completion task. This task type has been kept because it enables candidates to demonstrate their ability to listen for specific words or phrases and to produce written answers in response to the sentences. As the only productive task in the paper (one where candidates are required to write a response), this task replicates, to some extent, the real-world academic task of note-taking. Keeping Part 2 the same ensures that a wide range of test focuses, task types and text types is maintained across the paper as a whole. In

particular, the task balances and complements the new focus on interacting speakers in Part 3.

Part 3 predominantly moves from an interview with one speaker dominating to a task in which mainly two speakers take part in an equal two-way dialogue. There is no change to the testing of opinion, gist, detail and inference. However, it does mean that Part 3 may also include the additional testing focuses of attitude and agreement/disagreement across speaker turns. These testing focuses correspond to criterial features of C-level language users (Green 2012); in cognitive terms they tap into the higher-level processes involved in producing a discourse representation, which forms a key component of high-level listening performance (Field forthcoming 2013). The format of Part 3 enables the testing of more realistic, genuine interaction as both speakers take equal, shorter turns. As stated above, this change to Part 3 provides more balanced coverage given the change in Part 4, where a dialogue has been replaced by a series of monologues.

A new multiple-matching task has replaced the previous Part 4 task. The introduction of this task tests complex listening operations and allows for more testing focuses to be included in the test. This versatile task in conjunction with the revised format of Part 3 makes it possible to broaden the range of construct coverage, in particular testing interpreting context. Since the new Part 3 tests agreement/disagreement, which is the main focus of the old Part 4, the introduction of the multiple-matching task helps maintain the balance of testing focuses whilst widening the overall range of testing focuses.

### Overall

Although two items have been added to the paper, the overall timing has not changed. The new Listening paper has been able to maintain a wide range of test focuses by introducing the new Part 4 task and changing the interaction patterns found in Part 3. Trialling suggests that stakeholders view the changes positively.

**Table 4: Summary of the Speaking paper**

Paper/timing	Content
<b>Speaking</b> 16 mins (approx.)	<p><b>Part 1</b> A short conversation between the interlocutor and each candidate in which the candidates show their ability to use general interactional and social language.</p> <p><b>Part 2</b> A two-way conversation between the candidates. The candidates are given instructions with written and visual stimuli, which are used in a decision-making task. This task focuses on sustaining an interaction, exchanging ideas, expressing and justifying opinions, agreeing and/or disagreeing, suggesting, speculating, evaluating, reaching a decision through negotiation, etc.</p> <p><b>Part 3</b> An individual long turn by each candidate, followed by a discussion on topics related to the long turns. Each candidate in turn is given a written question to respond to. The focus is on organising a larger unit of discourse, expressing and justifying opinions, developing topics.</p>

## The Speaking paper

The speaking construct underlying Cambridge English exams emphasises production and interaction (Weir and Taylor 2011). Validation exercises which included matching the Cambridge English speaking specifications against an external model of cognitive processes were undertaken by Field (2011) as part of the *Studies in Language Testing* construct series *Examining Speaking*. Weir and Taylor (2011:302) conclude that the speaking specifications of the Cambridge English exams 'correspond closely to what we know of the cognitive processes involved in the production of speech'. Consequently, minor adjustments were made to the Speaking paper including: the removal of the set of discussion questions from Part 1 which were contained in Part 3, and Part 3 has been amended to direct a follow-up question to the listening candidate after each long turn to ensure an equal balance of talking time for each candidate. This replaces the open follow-up question, which overlaps with the wider discussion in the second phase of Part 3. Table 4 provides a summary of the new Speaking paper and Appendix 5 includes more detail of the structure of the paper.

The timing of speaking tasks is a key task parameter and it is important for test developers to provide adequate task response time which would give test takers sufficient time to address the requirements of the task (Galaczi and ffrench 2011, Weir 2005). The timing of Part 1 has been reduced from 3 minutes to 2 minutes by reducing the number of questions each candidate is asked. Now, after introductions, each candidate is asked only one question in this part. Speaking Examiners were consulted and it was felt that asking three questions to each candidate in the time allowed often meant that the candidates had to be cut short in their responses. The purpose of Part 1 is to 'warm up' the candidates and allow assessment of general interactional and social language. Asking each candidate just one question (after introductions) fulfils this purpose and allows for adequate assessment by giving them the opportunity to answer the question slightly more fully. This reduced task length is still in line with the progression across the Cambridge English tests in terms of time allocations by task (Galaczi and ffrench 2011).

Similarly, Part 3 has been reduced from 12 minutes to 10 minutes. This has been achieved by removing the first follow-up question. After the end of the 2-minute long turn the listening candidate was asked a reaction question, e.g. 'What do you think?', followed by a second question addressed to both candidates. While the first question brought the listening candidate back into the interaction, experience showed that it was not very generative. Now a single follow-up question is asked which extends the topic of the long turn; the question is asked first to the listening candidate and then the speaking candidate is invited to join in. Trialling showed that modifying the task in this way allowed candidates to produce a better breadth and depth of language in this part, i.e. the task better allowed the candidates the opportunity to produce adequate samples of language at C2 level. As noted earlier with the change to Part 1, the change in task length to Part 3 is in line with the progression in timing observed across the Cambridge English General English Speaking tests at different proficiency levels (Galaczi and ffrench 2011).

## Overall

Trialling of the new format took place in both the UK and overseas. The changes to Part 1 and Part 3 have made the Speaking paper slightly shorter, while preserving its aim of giving candidates every opportunity to produce C2-level language within a paired face-to-face format.

## Conclusion

Revising an exam presents a large number of challenges and potential pitfalls to the development team: the need to maintain the standard of the exam in terms of comparability of coverage and level with previous versions of the exam, while at the same time ensuring the revised version is fully fit for purpose in attempting to meet the needs of both existing and new target users. For the new *Cambridge English: Proficiency* this meant shortening the exam to improve the candidate experience by making it less of an 'endurance test' by reducing areas of overlap, without sacrificing language or skills coverage, and while also enhancing the exam's fitness for use for high-level academic study.

Has this project been successful in achieving its aims? All of the indications to date are that it has. Although live test performance data is not yet available, a large volume of pretesting data (live material which is trialled on a representative sample of the live test population) suggests that the level of the exam has been maintained. As described above, work by the Research and Validation team suggests that construct coverage and test reliability have been maintained. User feedback has been particularly positive as well. Elsewhere in this issue, Docherty and Howden's article on the results of a recent survey of teachers whose students have taken part in pretesting the new version shows an overwhelmingly positive attitude to the revised exam, which is certainly reassuring!

All aspects of the exam and its performance will of course be monitored very carefully once it goes live; however, the development team is confident that the new *Cambridge English: Proficiency* exemplifies how Cambridge English Language Assessment has balanced continuity and change.

## References

- Alderson, J C (2000) *Assessing Reading*, Cambridge: Cambridge University Press.
- Alderson, J C, Clapham, C and Wall, D (1995) *Language Test Construction and Evaluation*, Cambridge: Cambridge University Press.
- Bereiter, C and Scardamalia, M (1987) *The Psychology of Written Composition*, Hillsdale, NJ: Lawrence Erlbaum.
- Cambridge English Language Assessment (2013) *Principles of Good Practice: Quality Management and Validation in Language Assessment* available online: [www.cambridgeenglish.org/research-and-validation/quality-and-accountability](http://www.cambridgeenglish.org/research-and-validation/quality-and-accountability)
- Cambridge ESOL (2006) *Item Writer Guidelines: CPE Paper 3 Use of English*, Cambridge: Cambridge ESOL internal report.
- Cambridge ESOL (2008) *Certificate of Proficiency in English. Handbook for Teachers*, Cambridge: Cambridge ESOL.
- Cambridge ESOL (2011) *Revised CPE Specification Summary*, Cambridge: Cambridge ESOL internal report.

- Cambridge ESOL (2012) *Cambridge English: Proficiency. Handbook for Teachers*, available online: [www.teachers.cambridgeenglish.org/ts/exams/academicandprofessional/cpe](http://www.teachers.cambridgeenglish.org/ts/exams/academicandprofessional/cpe)
- Carson, J and Leki, I (Eds) (1993) *Reading in the Composition Classroom: Second Language Perspectives*, Boston: Heinle and Heinle.
- Council of Europe (2001) *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*, Cambridge: Cambridge University Press.
- Field (2011) Cognitive validity, in Taylor, L (Ed.) *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, Studies in Language Testing volume 30, Cambridge: UCLES/Cambridge University Press, 65-111.
- Field (forthcoming 2013) Cognitive validity, in Geranpayeh, A and Taylor, L (Eds) *Examining Listening: Research and Practice in Assessing Second Language Listening*, Studies in Language Testing volume 35, Cambridge: UCLES/Cambridge University Press, 77-151.
- Fried-Booth, D (2004) Set texts in CPE Writing, *Research Notes* 18, 12-17.
- Galaczi, E and French, A (2011) Context validity, in Taylor, L (Ed.) *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, Studies in Language Testing volume 30, Cambridge: UCLES/Cambridge University Press, 112-170.
- Green, A (2012) *Language Functions Revisited. Theoretical and Empirical Bases for Language Construct Definition Across the Ability Range*, English Profile Studies volume 2, Cambridge: UCLES/Cambridge University Press.
- Khalifa, H and Weir, C J (2009) *Examining Reading: Research and Practice in Assessing Second Language Reading*, Studies in Language Testing volume 29, Cambridge: UCLES/Cambridge University Press.
- Lim, G S (2012) Developing and validating a mark scheme for Writing, *Research Notes* 49, 6-10.
- North, B (2004) Relating assessments, examinations, and courses to the CEF, in Morrow, K (Ed), *Insights from the Common European Framework*, Oxford: Oxford University Press, 77-90.
- Read, J (2000) *Assessing Vocabulary*, Cambridge: Cambridge University Press.
- Saville, N (2003) The process of test development and revision within UCLES EFL, in Weir, C J and Milanovic, M (Eds) *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913-2002*, Studies in Language Testing volume 15, Cambridge: UCLES/Cambridge University Press, 57-120.
- Shaw, S D and Weir, C J (2007) *Examining Writing: Research and Practice in Assessing Second Language Writing*, Studies in Language Testing volume 26, Cambridge: UCLES/Cambridge University Press.
- Spillett, H (2012) The revision of the Cambridge English: Proficiency Writing paper, *Research Notes* 49, 2-5.
- Weir, C J (2003) A survey of the history of the Certificate of Proficiency in English (CPE) in the twentieth century, in Weir, C J and Milanovic, M (Eds) *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913-2002*, Studies in Language Testing volume 15, Cambridge: UCLES/Cambridge University Press, 1-56.
- Weir, C J (2005) *Language Testing and Validation: An Evidence Based Approach*, Basingstoke: Palgrave Macmillan.
- Weir, C J and Jin, Y (1996) *Theory of Reading*, unpublished mimeo, CALS, University of Reading.
- Weir, C J and Milanovic, M (2003) (Eds) *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913-2002*, Studies in Language Testing volume 15, Cambridge: UCLES/Cambridge University Press.
- Weir C J and Porter D (1994) The multi-divisible or unitary nature of reading: The language tester between Scylla and Charybdis, *Reading in a Foreign Language* 10 (2), 1-19.
- Weir, C J and Taylor, L (2011) Conclusions and recommendations, in Taylor, L (Ed.) *Examining Speaking: Research and Practice in Assessing Second Language Speaking*, Studies in Language Testing volume 30, Cambridge: UCLES/Cambridge University Press, 293-313.
- Weir, C J, Vidaković, I and Galaczi, E D (forthcoming 2013) *Measured Constructs: A History of Cambridge English Language Examinations 1913-2012*. Studies in Language Testing volume 37, Cambridge: UCLES/Cambridge University Press.

## Appendix 1: Summary of the new and old *Cambridge English: Proficiency exam*

New exam		Old exam	
Paper/timing	Content	Paper/timing	Content
<b>Reading and Use of English</b> 1 hr 30 mins	7 parts 53 questions	<b>Reading</b> 1 hr 30 mins	4 parts 40 questions
<b>Writing</b> 1 hr 30 mins	2 parts: 1 compulsory question; 1 from a choice of 5 (including the set text options)	<b>Writing</b> 2 hrs	2 parts: 1 compulsory question; 1 from a choice of 5 (including the set text options)
<b>Listening</b> 40 mins (approx.)	4 parts 30 questions	<b>Use of English</b> 1 hr 30 mins	5 parts 43 questions
<b>Speaking</b> 16 mins (approx.)	3 parts: interview; collaborative task; individual long turns and follow-up discussion	<b>Listening</b> 40 mins (approx.)	4 parts 28 questions
		<b>Speaking</b> 19 mins (approx.)	3 parts: interview; collaborative task; individual long turns and follow-up discussion
<b>Total timing: 3 hours 56 minutes</b>		<b>Total timing: 5 hours 59 minutes</b>	

## Appendix 2: Summary of the Reading and Use of English paper

<b>Word count</b>	2,900–3,400
<b>Timing</b>	1 hr 30 mins
<b>Number of parts</b>	7
<b>Number of questions</b>	53
<b>Marks</b>	For Parts 1–3, each correct answer receives 1 mark; for Part 4, each correct answer receives up to 2 marks; for Parts 5–6, each correct answer receives 2 marks; for Part 7, each correct answer receives 1 mark. There are a total of 72 marks available for the test.

### Structure and tasks of the Reading and Use of English paper

Part	Task type	Focus	Format	Number of questions
1	Multiple-choice cloze	The main focus is on vocabulary, e.g. idioms, collocations, fixed phrases, complementation, phrasal verbs, semantic precision.	A single text with eight gaps. Candidates must choose one word or phrase from a set of four to fill each gap.	8
2	Open cloze	The main focus is on awareness and control of grammar with some focus on vocabulary.	A modified cloze test consisting of a text with eight gaps. Candidates think of the word which best fits each gap.	8
3	Word formation	The main focus is on vocabulary, in particular the use of affixation, internal changes and compounding in word formation.	A text containing eight gaps. Each gap corresponds to a word. The stems of the missing words are given beside the text and must be changed to form the missing word.	8
4	Key word transformations	The focus is on grammar, vocabulary and collocation.	Six discrete items with a lead-in sentence and a gapped response to complete in three to eight words including a given 'key' word.	6
5	Multiple choice	The focus is on understanding of detail, opinion, attitude, tone, purpose, main idea, implication, text organisation features (exemplification, reference).	A text followed by 4-option multiple-choice questions.	6
6	Gapped text	The focus is on understanding of cohesion, coherence, text structure, global meaning.	A text from which paragraphs have been removed and placed in jumbled order after the text. Candidates must decide from where in the text the paragraphs have been removed.	7
7	Multiple matching	The focus is on understanding of detail, opinion, attitude, specific information.	A text, or several short texts, preceded by multiple-matching questions. Candidates must match a prompt to elements in the text.	10

## Appendix 3: Summary of the Writing paper

<b>Timing</b>	1 hr 30 mins
<b>Guideline length</b>	<b>Part 1:</b> 240–280 words <b>Part 2:</b> 280–320 words
<b>Marks</b>	Each task carries equal marks

### Structure and tasks of the Writing paper

Part	Task type and focus	Format
1	An essay with a discursive focus.	Candidates are required to write an essay summarising and evaluating the key ideas contained in two input texts of approximately 100 words each. These texts may contain complementary or contrasting opinions drawn from a variety of authentic, contemporary sources.
2	A contextualised task from a range of text types, including questions on set texts.	Candidates have a choice of task in Questions 2–5. Each task provides candidates with a clear context, topic, reason for writing and target reader. Questions 5a and 5b offer a choice of two tasks based on the set texts. The output text types for Questions 2–4 are: article, letter, report, review; for Questions 5a and 5b there is also the additional text type of an essay. Part 2 will always offer a range of text types, but will not necessarily include all the text types on any one paper.

## Appendix 4: Summary of the Listening paper

<b>Timing</b>	Approximately 40 mins
<b>Number of parts</b>	4
<b>Number of questions</b>	30
<b>Recording information</b>	The instructions for each task are given on the question paper and are also heard on the recording. These instructions include the announcement of pauses of specified lengths, during which candidates can familiarise themselves with the task. A variety of voices, styles of delivery and accents will be heard in each paper to reflect the various contexts presented in the recordings, as appropriate to the international context of the test takers.
<b>Marks</b>	Each correct answer receives 1 mark

### Structure and tasks of the Listening paper

Part	Task type	Focus	Format	Number of questions
1	Three-option multiple choice	The focus is on gist, detail, function, purpose, topic, feeling, attitude, opinion and agreement/disagreement.	Three short unrelated texts lasting approximately 1 min each, consisting of either monologues or exchanges between interacting speakers. There are two multiple-choice questions per text, with three options.	6
2	Sentence completion	The focus is specific information and stated opinion.	A monologue lasting 3 to 4 mins.	9
3	4-option multiple choice	This focuses on opinion, gist, detail, inference, attitude, agreement/disagreement.	A text involving interacting speakers lasting 3 to 4 mins.	5
4	Multiple matching	The focus here is gist, attitude, main points and interpreting context.	Five short themed monologues of approximately 35 seconds each. There are two tasks. Each multiple-matching task requires selection of the five correct options from a list of eight.	10

## Appendix 5: Summary of the Speaking paper

<b>Timing</b>	16 mins
<b>Interaction pattern</b>	Two candidates and two examiners. One examiner acts as both interlocutor and assessor and manages the interaction either by asking questions or setting up the tasks for candidates. The other acts as assessor and does not join in the conversation.
<b>Marks</b>	Candidates are assessed on their performance throughout the test.

### Structure and tasks of the Speaking paper

Part	Task type and format	Focus	Timing
1	A short conversation between the interlocutor and each candidate.	Candidates show ability to use general interactional and social language.	2 mins (3 mins for groups of 3)
2	A two-way conversation between the candidates. The candidates are given instructions with visual stimuli which they use in a decision-making task.	The focus is on sustaining an interaction, exchanging ideas, expressing and justifying opinions, agreeing and/or disagreeing, suggesting, speculating, evaluating, reaching a decision through negotiation, etc.	4 mins (6 mins for groups of 3)
3	An individual long turn by each candidate, followed by a brief discussion on the topic related to the long turns. Each candidate in turn is given a written question to respond to. The interlocutor leads a discussion to explore further the topic covered in the individual long turns.	The focus is on organising a larger unit of discourse, expressing and justifying opinions, developing topics etc.	10 mins (15 mins for groups of 3)

# A look into the future

**MICHAEL MILANOVIC** CHIEF EXECUTIVE, CAMBRIDGE ENGLISH LANGUAGE ASSESSMENT

## Introduction

The role of English as a *lingua franca* has evolved rapidly over the last 100 years. English now facilitates international communication and migration, its uses and levels described and validated in terms of internationally developed and accepted frameworks. English teaching, often formatively linked with assessment, is increasingly aimed at meeting the communication needs of those studying through the medium of English in schools, the workplace and academic contexts.

Over the last few years, Cambridge English Language Assessment has continually responded to this changing world and, as we look to the future, there is no reason to assume this will not continue in the next 100 years. The ability to manage change and innovate in creative ways while holding true to core values and principles has been a strength in the past which needs to be built on in the future. Therefore, in this article I will consider the way things might be for Cambridge English Language Assessment in the future by picking up on two central themes from the past that have underpinned the success so far: managing change while maintaining continuity and achieving innovation; and the widening participation of stakeholders within the expanding Cambridge English Language Assessment community.

## Managing change in English language learning and teaching

### The changing context of learning

We have witnessed an exponential growth in candidature over the last 100 years from the three 1913 CPE candidates to the more than 4 million candidates in 2013, who will be taking a Cambridge English exam selected from a comprehensive range of assessment products designed for different levels and purposes. We are clearly seeing the role of English evolve as parents, schools, employers, local, regional and national governments recognise the competitive advantages a good command of English can bring personally and more broadly at a societal level.

We have moved from a world where English is just another modern language normally taught in secondary school or in private language schools to one where it is a core subject in the curriculum of most countries, often taught from the early years of primary education. This change has led to an intensification of English language instruction, in both private and state-funded institutions, as witnessed by the increasing hours spent on English instruction as well as the growth of Content and Language Integrated Learning (CLIL). As a result, not only will the demand for learning materials and exams suitable for these learners continue to increase in the future, but so too will the need for localised content that is closely referenced to international standards. The situation where international English exams are the same for

everyone is beginning to change. In such a context Cambridge English Language Assessment is well placed to link local standards to international ones. Projects are now underway, in collaboration with Cambridge University Press, where local teaching and learning materials will be complemented by assessment materials focused on particular contexts of use.

A consequence of the increased role of English in the classroom is that the levels achieved by learners in compulsory education keep on rising with many countries now targeting B1 and sometimes B2 as an exit level. Successful adult learners will need to focus on specific work-related skills in future to a much greater extent than before. We have seen the development of exams over the last few years to meet these demands such as *BULATS (Business Language Testing Service)*, *Cambridge English: Business (BEC)*, *Cambridge English: Legal (ILEC)*, *Cambridge English: Financial (ICFE)* as well as the inclusion of medical English tests in the Cambridge English suite of exams. This trend towards customisation in both learning and assessment at the higher levels of proficiency is set to continue.

### A global community of English teachers

Although there has been a focus on developing the skills of teachers of English since 1913, the last two or three decades have seen that focus gain in importance as the demand for qualified English teachers grows to meet language learning needs. Cambridge English qualifications for teachers are already very widely taken but their use will become even more significant in the future. In conjunction with recent developments such as the Cambridge English Teacher website we see Cambridge set to play a key role in the upskilling of English teachers around the world for decades to come. Such upskilling may come through international qualifications and ones tailored to local requirements; face-to-face and online training courses; the intensification of professional guidance and exchange through online and webinar contact with teachers, and ultimately the creation of a global community of English language teachers all interacting with Cambridge English Language Assessment.

### Impact of technology on assessment

We have maintained a strong awareness of the impact of technology on the administration and delivery of examinations, from the early adoption of basic computing for aspects of exams processing in the 1950s to word processing in the 1980s and local area networks and personal computers in the 1990s. The importance of computer-based testing has grown and developed over the last 20 years and the internet has allowed for much more extensive interaction with centres, schools, candidates, teachers, parents and a full range of other stakeholders over the last decade.

With the ever-increasing number of developments on this front, it is fully predictable that changes in computer-based testing and computerised exam management systems will

pave the way to being able, for example, to offer even more exam sessions a year. This will be facilitated by various types of computer-assisted marking and automated test construction systems. Within the next few years we will expect to see such systems come into their own, allowing test centres to schedule examinations that are entirely convenient to candidates. This already happens with *BULATS* and the *Cambridge English Placement Test*. Such systems offer adaptive assessment based on large banks of material and are extremely valuable in certain contexts. However, they are not necessarily as in depth, from a pedagogical point of view, as the range of linear Cambridge English examinations, and it is these that we will also see being offered on demand in the years to come.

We have spent decades building extensive banks of test materials that supply both linear and adaptive testing systems, all of which are calibrated to a common scale and described by a comprehensive attribute system. We expect that teachers will soon be able to construct diagnostic tests from sections of these banks that will allow for much more fine-tuned assessment than even the level-based linear tests can currently provide. This will allow more customisation of content to local contexts and is in line with the broadening mission of Cambridge English Language Assessment to engage more with the entirety of language education. For example, in collaboration with Cambridge University Press, we are producing a new multimedia course based on the new and strengthened formative assessment model now referred to as *learning-oriented assessment* (LOA). Integrating learning and assessment by providing teachers and learners with timely and relevant feedback which can be used to tailor learning is key to LOA and we need to be leaders in this field.

There is little doubt that the future of Cambridge English Language Assessment will depend on how well technology is incorporated into its work to deliver benefits for its users. No less so, of course, as assessment becomes integrated into learning. In this context, learners, whether in classrooms or elsewhere, are more likely to be tested when ready and we will need technology to enable us to do this successfully. The challenge of test security, including the prevention of the fraudulent use of test results will remain a key focus of attention, but our research capability, our practical expertise and our deployment of technology coupled with the vigilance of our global network of centres will enable us to combat these threats.

However, ensuring the *usefulness* of tests and assessment more generally through the deployment of technology is a challenge, and we must try to take advantage of the benefits technology has to offer without the technology tail wagging the learning and assessment dog. The question we need to continually ask is will technology take us forward or will it impose an approach, which while appealing on some levels, forces a backward step in pedagogical terms? This problem has been faced many times over the last hundred years and is unlikely to disappear.

### The research agenda

Just as the future of a major assessment organisation such as Cambridge English Language Assessment must involve major efforts to keep abreast with developments in computer

science and technology, so also its future depends on its awareness and use of insights from linguistics, theoretical and applied. The research thrust into aspects of the validity, reliability, impact and practicality of our exams has been a major research theme for the organisation since the 1980s. We must continue to participate in research that will contribute to an eventual model of learning, which will allow us to produce tailored, diagnostic feedback and better targeted support for learners. Test takers increasingly require to know, in valid and reliable terms, what they should expect to be able to do at each stage of their learning. The assessments they use must provide finer-grained information with respect to features clearly specified as *critical* to levels of proficiency in the language, in terms of pronunciation, grammar, vocabulary, language notions and functions, communicative language skills and so on. It is precisely in this area that the English Profile Programme is playing a key role. It has been an important driver for *criticality* research over the last decade and this research will continue to grow in relevance and importance. In addition, our role in providing accurate reference to the Common European Framework of Reference for languages (CEFR) will also prove invaluable as the use of this framework gathers momentum. Schools, colleges, local and national authorities as well as employers will need to measure progress in a wide range of contexts reliably against international benchmarks and we are ideally placed to help. Ultimately, by linking materials empirically to an ability continuum, learners and teachers will be able to interact much more meaningfully with the level targeted by these materials.

Strengthening our research focus is the fact that Cambridge English Language Assessment is a part of one of the world's leading universities and it is centrally involved in a programme of language research across the University through the University's Language Sciences Research Initiative, established to promote interaction across different departments. Five key research themes have been identified for the Initiative: language communication and comprehension; language learning across the lifespan; language change and diversity; human language technologies, and Cambridge English Language Assessment. This latter theme will continue to cover the kind of interdisciplinary research projects shared with Cambridge University Press and University departments such as the Department for Theoretical and Applied Linguistics (DTAL) the Computer Laboratory and the Speech Group within the Engineering Department.

Particular areas of mutual interest are likely to be the ways in which researchers in Cambridge have been analysing learner data, including corpora, to understand progression across levels of proficiency, and the development of automated systems for the assessment of English speech and writing based on computational techniques derived from natural language processing. For instance, following the example of the successful Cambridge Learner Corpus, Cambridge is rightly beginning to build spoken English corpus systems and automated sampling of quite a proportion of its speaking exam data. Cambridge English Language Assessment will continue to learn from and contribute to the Language Sciences Research Initiative, thus bringing

together this broad-ranging research initiative for very real educationally and socially valuable outcomes.

### **Building partnerships**

One hundred years of experience in a radically changing world has shown the importance of building relationships and valuing partnerships. We have built strong collaborative relationships all over the world, for example with the British Council, IDP (International Development Program) in Australia, the Cultura network in Latin America, the National Education Examinations Authority (NEEA) in China, the Language Training and Testing Center (LTTC) in Taiwan, STAR TV in India, the French Ministry of Education and FERE Madrid (a federation of Spanish schools), to name but a few. A collaboration with the Italian Ministry of Education, for example, is an important indicator of the role Cambridge English Language Assessment will continue to play in the future. Our Admissions Testing Service is currently assessing students wishing to study medicine in Italy through the

medium of English. These partnerships provide an opportunity for an interaction between international standards and local requirements that is set to become so important in the years to come. As a result, we will see more assessment staff trained in Cambridge based in country and regional offices around the world. They will be well placed to help bind the Cambridge English Language Assessment values, processes and procedures with local ones. Our people, at home and abroad, will remain at the heart of what we do – an ethos marked at the recent Cambridge English Centenary Launch event, pictured below.

We will continue in the next 100 years to set great store in developing appropriate testing materials, researching them effectively, making sure they cover the areas that need to be assessed, providing evidence to back that up and continuing to improve those materials on the basis of the evidence that has been gathered.



**Cambridge English Language Assessment staff and consultants marking the centenary at an event in the Guildhall, Cambridge, 28 January 2013**

# 2012 Research Notes reader survey

**FIONA BARKER** RESEARCH AND VALIDATION GROUP, CAMBRIDGE ENGLISH LANGUAGE ASSESSMENT

**COREEN DOCHERTY** RESEARCH AND VALIDATION GROUP, CAMBRIDGE ENGLISH LANGUAGE ASSESSMENT

## Introduction

The first issue of *Research Notes* was published in March 2000 describing itself as a 'newsletter on current developments in the research, validation and test development work carried out by the EFL Division at UCLES' (University of Cambridge Local Examinations Syndicate 2000:1). Since then, over 350 articles have been published in it on a variety of topics related to language learning and assessment. It has evolved over the years from a short 'newsletter' into a respected publication increasingly being referenced in other publications.

*Research Notes* is currently mailed to readers in 107 different countries with the top 10 countries, in terms of the number of issues distributed, being the United Kingdom, France, China, India, Argentina, Spain, Australia, Italy, Greece and Brazil. In order to determine whether we are meeting our readers' expectations and needs as we approached our 50th issue, the editors included a Reader Survey in issue 49 (August 2012). This article summarises the results of the survey and describes the readership who took part.

## The survey

The Reader Survey went live at the beginning of August 2012 and in the following two months, almost 300 individuals provided their views on the content and approach of *Research Notes* together with how they accessed and used this publication.

The survey was delivered via an online survey tool, in hard copy in *Research Notes* issue 49 (August 2012) and it was also publicised in the August issue of the Cambridge English Teacher Support ezine ([www.teachers.cambridgeenglish.org/ts/emedial](http://www.teachers.cambridgeenglish.org/ts/emedial)) to ensure that both current readers and prospective readers had the opportunity to respond. We received 289 responses with 75% responding to our email invitation (institutions and individuals on our mailing list who receive hard copies every quarter) and 25% of responses came from those who accessed the website link.

We asked 13 questions plus provided an open-ended response for comments, and the majority of respondents completed the whole survey. We report below on the characteristics of the readership, what they are most interested in reading about and finally on their reading behaviour.

## Reader characteristics

We were interested in finding out about the main occupation of our readers, as shown in Table 1.

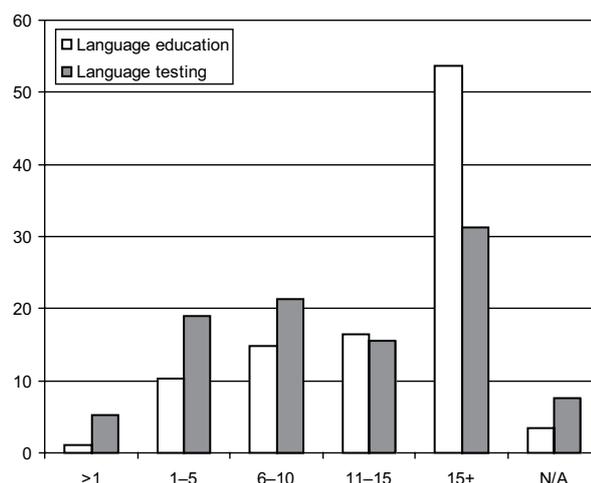
**Table 1: Main occupation of respondents (%)**

Occupation	Percentage
Teacher	43
Lecturer	18
Administrator	10
Language testing practitioner	9
Researcher	8
Student	6
Other	6

The 'other' category includes occupations both within education (e.g. item writers, curriculum developers, consultants, etc.) and outside education (e.g. a member of the police force, a journalist, etc.). Unsurprisingly, the highest proportion of respondents were teachers, followed by lecturers and administrators, which demonstrates the wide readership of *Research Notes* in terms of both practitioners and researchers.

In response to the question 'How long have you been involved in language education/language testing?', the majority of respondents for both questions indicated that they had more than 15 years' experience as seen in Figure 1. While 70% of respondents have been involved in language education for 11 or more years, only 45% have been involved in language testing for a similar length of time.

**Figure 1: Number of years' experience in language education/testing (%)**



## Topics of interest

When asked: 'Which exam domains are you most interested in reading about?', respondents indicated whether they were very/somewhat/not very interested in six options. The responses for 'very interested in' are summarised in Table 2.

**Table 2: Exams of most interest to respondents (%)**

Exam domains	Very interested in (%)
Academic and Professional exams	79
General English exams	75
Cambridge English teacher qualifications	59
Young Learners and for Schools exams	49
Business English exams	43

Table 2 shows that the majority of respondents are most interested in Academic and Professional exams followed by General English exams. In response to the 'other' option, readers specified particular exams such as *IELTS* (which is covered under the Academic and Professional English exam category), exams for specific professions such as for health professionals, engineers or architects, exams related to Content and Language Integrated Learning (CLIL) and exams in languages other than English.<sup>1</sup>

The next question investigated the skills/systems readers are most interested in. Table 3 shows that the productive skills were the most popular options with listening and reading following close behind and then vocabulary and grammar. In response to this question, several people made use of the 'other' category to point out that fluency-related skills such as discourse management, phonology and in particular pronunciation were missing from the list of skills provided. Intercultural communicative competence and lexical chunks were also specified in response to the 'other' option.

**Table 3: Skills and systems of most interest to respondents (%)**

Skills/systems	Very interested in (%)
Speaking	85
Writing	83
Listening	75
Reading	72
Vocabulary	68
Grammar	58

**Table 4: Topics of most interest to respondents (%)**

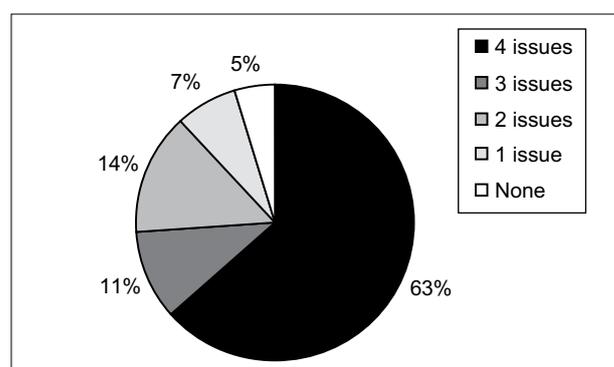
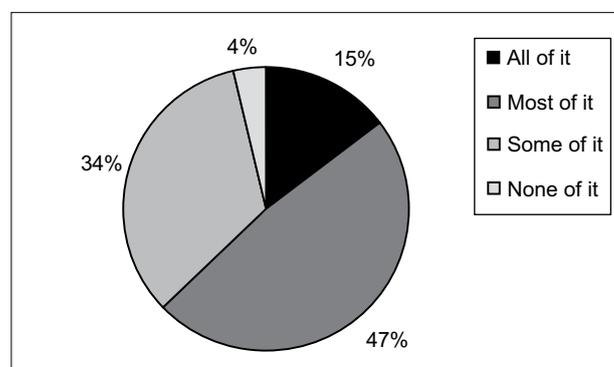
Topics	Very interested in (%)
Teaching/learning	83
Test design/development	76
Learning-oriented assessment/formative assessment	66
Validity/reliability	61
Research methods	58
Common European Framework of Reference (CEFR)	55
Standard setting/benchmarking	55
Washback/impact	52
Technological developments in testing	50
Operational processes of an exam board	47
English Profile	46
Content and Language Integrated Learning (CLIL)	46
Conference/publication updates	46
Using corpora	42

In relation to the topics that readers would like to see included in *Research Notes*, Table 4 summarises the responses for 'very interested in'.

Teaching/learning was not surprisingly of most interest to respondents given the large proportion of teachers who completed the survey.

## Reader behaviour

Fifty-four percent of readers first found out about *Research Notes* from the Cambridge English Language Assessment website, 19% from a direct mailing and 16% from an instructor or colleague. It is encouraging to note that 63% of respondents read every issue per year while almost the same percentage read all or most of each issue (see Figures 2 and 3 respectively).

**Figure 2: Number of issues read per year in whole or in part****Figure 3: The amount of each issue read by respondents**

Over half of respondents (57%) have accessed further information from our website after reading *Research Notes* and 82% found it easy to locate the information that they were looking for when they visited the website. The most frequently cited information accessed by respondents is found in Table 5.

**Table 5: Information accessed from the website**

Information accessed	Percentage
Publications (includes back issues of <i>Research Notes</i> , <i>Studies in Language Testing</i> volumes, references, <i>IELTS Research Reports</i> , etc.)	29
Assessment product information	28
Teaching resources	11
Exam statistics	9

<sup>1</sup> For information about the range of exams and teaching qualifications provided by Cambridge English, see: [www.cambridgeenglish.org/exams/](http://www.cambridgeenglish.org/exams/)

In response to the open-ended final question, 'Do you have any other comments on *Research Notes*?', 90 people left comments (32% of the respondents) and it is encouraging to note that most of these responses were positive. Some examples of typical comments are presented below:

*I find the research notes really valuable, and recommend them to both coursework and research students. Thank you for making them so accessible, and free!*

*It's a significant professional publication and it has kept me up to date. In fact, I have insisted that each of my research scholars also reads and uses it in their studies.*

*I sometimes feel isolated from mainstream publications (often for financial reasons), therefore I feel content with Cambridge ESOL and the RN I receive on a regular basis.*

*I got my teaching degree almost 30 years ago, if it weren't for Research Notes that keep me updated I wouldn't be able to keep teaching and preparing my students for exams!*

*An extremely useful publication which has come to the rescue several times – while doing my PhD, teaching Delta candidates and applying for funding. The paper copies are brill [brilliant] as I can read them on the train, scribble in the margins etc.*

*They are most beneficial to cover practical problems that journals are not usually interested in.*

*Thanks for offering and sending the RN and for opportunity to respond in a survey such as this. I feel RN does scratch a need/fill a niche.*

The comments above highlight three recurring points: *Research Notes* is regarded not only as a valuable academic publication but also as an important source for practical assessment-related information which is not often published elsewhere. It is used by a wide range of practitioners (from teachers to academics) and it fills a void for readers with less access to academic journals or materials.

Finally, a few individuals inquired about submission procedures. Although we appreciate such interest, at present the primary role of the publication is to inform readers about our work and our research programme. This allows us to remain a free publication which many respondents pointed out was of particular value in their contexts where access to research may be limited. However, thought is currently being given to widening the scope of *Research Notes* in the near future, which may include calls for submissions.

## Conclusion

We intend to continue publishing *Research Notes* on a quarterly basis, as now, and aim to maintain its ethos and interest to a wide readership which is representative of those involved in teaching and assessment around the world. Clearly our main focus will remain reporting on our key research and validation activities, however we hope to be able to meet the needs expressed by the respondents to this survey in future issues.

We would like to thank all those who took part and if you have any further suggestions for improving *Research Notes*, or would like to join the mailing list or update your contact details please email [validation@cambridgeenglish.org](mailto:validation@cambridgeenglish.org) with the subject: Research Notes. We look forward to another successful 50 issues.

## References

University of Cambridge Local Examinations Syndicate (2000) *Research Notes 1*, available online: [www.cambridgeesol.org/rs\\_notes/rs\\_nts1.pdf](http://www.cambridgeesol.org/rs_notes/rs_nts1.pdf)

# Caroline Clapham IELTS Masters Award 2012

For over a decade now, the *IELTS* partners have presented the Caroline Clapham *IELTS* Masters Award annually to the Master's-level dissertation or thesis in English which makes the most significant contribution to the field of language testing.

Recently, the *IELTS* Research Committee announced the selection of Veronika Timpe as the winner of the 2012 award. Her dissertation investigated the testing of sociopragmatic strategies, an under-researched area in language testing. One reviewer wrote that Ms Timpe 'is to be congratulated on identifying a very significant gap in the research . . . and identifying empirical support for her findings . . . [taking] research in this area beyond what has been previously investigated'. Another reviewer commented that 'the author has demonstrated excellent research skills and knowledge,

and has conducted an ambitious project with rigour and precision'. The dissertation was supervised by Luke Harding, and was submitted to Lancaster University. The abstract for the dissertation appears below. In addition, the Committee issued a 'Highly Commended' certificate to Anne-France Pinget of Utrecht University for the excellent quality of her dissertation on fluency and accent in L2 speech.

Veronika will be presented with her award - a certificate and a cheque for £1,000 - at the 2013 Language Testing Research Colloquium in Seoul, Republic of Korea. Qualified individuals who would like to join the 2013 competition are invited to visit [http://ielts.org/researchers/grants\\_and\\_awards/ielts\\_masters\\_award.aspx](http://ielts.org/researchers/grants_and_awards/ielts_masters_award.aspx) for details of the competition and submission guidelines.

## Strategic Decoding of Sociopragmatic Utterances: A Think-aloud Validation Study

**Ms Veronika Timpe** Lancaster University

Sociopragmatics has proven to be a challenging field in language testing given that pragmatic expectations and assessments are highly culture- and context-specific (Liu 2007). Thus, it is challenging to avoid construct-irrelevant variance and to draw valid inferences on the basis of overall test scores. In an attempt to answer Roever's call for a 'broadening of the evidence base that allows extrapolation inferences to a target domain of social language use in academic and non-institutional contexts' (2011:3), this study investigated the cognitive processes of university-level German learners of English when solving receptive sociopragmatic assessment tasks.

Two groups of university-level EFL students with different amounts of exposure to the target language environment were asked to answer seven multiple choice discourse completion tasks taken from the American English Sociopragmatics Comprehension Test (AESCT), an intercultural sociopragmatics comprehension test that focuses on U.S.-

American English as well as the academic context in the United States. Verbal report methodology was used to access respondents' cognitive processes while they were working on the tasks.

By means of a grounded theory analysis, the author compiled a taxonomy of 24 strategies in three categories: recall, evaluation, and other. A contrastive between-group investigation showed that respondents with much more exposure to the target language context showed a greater ability to contextualise, while candidates without exposure revealed a stronger reliance on the text and evaluation strategies to compensate for the lack of (experiential) knowledge. Although a final analysis with regard to the substantive aspect of construct validity (Messick 1989, 1996) revealed that patterns in the data supported the trends hypothesised in the test construct, it also exposed some items that underrepresent the construct.

# Studies in Language Testing

Four new volumes from the *Studies in Language Testing* (SiLT) series will be in print at the beginning of 2013.

Volume 35 of the series, edited by Ardeshtir Geranpayeh and Lynda Taylor, is entitled *Examining Listening: Research and Practice in Assessing Second Language Listening*. Test developers need to provide a clear explication of the language ability constructs that underpin the tests they offer in the public domain. *Examining Listening* provides such an explication, through its focuses on the application of a socio-cognitive theoretical framework for validating tests of second language listening ability. The framework is applied to tasks in Cambridge English Listening tests from a number of different validity perspectives and the authors show how an understanding and analysis of the framework and its components can assist test developers to operationalise their listening tests more effectively, especially in relation to the key criteria that differentiate one proficiency level from another. As a companion volume to the previously published 'construct' volumes, *Examining Writing* (2007), *Examining Reading* (2009) and *Examining Speaking* (2011), this volume will be of considerable interest to test providers who wish to validate their own tests in a systematic and coherent manner, as well as to academic researchers and graduate students in the field of language assessment. *Examining Listening* will be published in March 2013.

Volume 36, entitled *Exploring Language Frameworks: Proceedings of the ALTE Kraków Conference, July 2011*, explores the impact of language frameworks on learning, teaching and assessment, viewed from the perspective of policies, procedures and challenges. The volume editors, Evelina D Galaczi and Cyril J Weir, bring together a selection of edited papers based on presentations given at the 4th International Conference of the Association of Language Testers in Europe (ALTE) held in Kraków, Poland, in July 2011. The selected papers focus on several core strands addressed during the conference, such as the role of language frameworks in migration and multilingual policy and practice, the use of frameworks in educational contexts, and practical issues associated with the application of frameworks. With its broad coverage of key issues and combination of theoretical insights and practical advice, this volume will be of considerable interest for academics, employers and policy-makers in

Europe and beyond. *Exploring Language Frameworks* was published in January 2013.

Two SiLT volumes which take a historical perspective will also be published as part of the series in early 2013. The volumes are designed as centenary volumes to mark 100 years of Cambridge English exams. Volume 37, entitled *Measured Constructs: A History of Cambridge English Language Examinations 1913–2012*, by Cyril J Weir, Ivana Vidaković and Evelina D Galaczi, focuses on how approaches to measuring English language ability have evolved worldwide and at Cambridge over the last 100 years. The volume takes the reader from the first form of the Certificate of Proficiency in English offered to three candidates in 1913, a serendipitous hybrid of legacies in language teaching from the previous century, up to the current Cambridge approach to language examinations, where the language construct to be measured is seen as the product of the interactions between a targeted cognitive ability, a highly specified context of use and a performance level based on explicit and appropriate criteria of description. *Measured Constructs* is a rich source of information on how changes in language pedagogy, together with wider socio-economic factors, have shaped the development of English language exams in Cambridge English Language Assessment and worldwide over the last century. As such, it will be of considerable interest to researchers, practitioners and graduate students in the field of language assessment. *Measured Constructs* will be published in April 2013.

Volume 38 of the series, entitled *Cambridge English Exams – The First Hundred Years: A History of English Language Assessment from the University of Cambridge, 1913–2013*, is authored by Roger Hawkey and Michael Milanovic and will be published in the coming months. The authors trace the history of the Cambridge English exams through their first 100 years, setting them in the context of wider educational and academic developments. They pay particular attention to the contribution of the dedicated individuals in Cambridge and around the world who have contributed to the success of the exams and to their positive educational impact. This volume will be of value to anyone interested in language teaching and assessment, applied linguistics or educational history, and to the thousands of people who are part of the wider Cambridge English Language Assessment 'family'.

# ALTE report

## ALTE Language Testing Courses, September 2012

ALTE (the Association of Language Testers in Europe) ran two one-week Language Testing Courses at Hughes Hall in Cambridge in September. **The ALTE Introductory Course in Language Testing** took place on 3-7 September and was taught by Dr Lynda Taylor, University of Bedfordshire and Consultant to Cambridge English Language Assessment, and Professor Cyril Weir, University of Bedfordshire. **The ALTE Course in Understanding the C-levels to Assess Language for the Professions** took place on 10-14 September and was taught by Dr Anthony Green, University of Bedfordshire and Dr Fiona Barker, Cambridge English Language Assessment.

Participants on the courses came from many countries: Austria, Belgium, China, the Czech Republic, Germany, Italy, Luxembourg, Portugal, Saudi Arabia, Spain, the UK and the United States, and feedback from both courses has been very positive. Next summer the courses will take place in Sofia, Bulgaria and will be hosted by the Department for Language Teaching and International Students (DLTIS) at Sofia University. Further details will be available on the ALTE website.

## ALTE 42nd Meeting and Conference

Over 100 delegates attended ALTE's 42nd bi-annual Meeting and Conference held in Munich on 21-23 November 2012 and hosted by the Goethe-Institut.

The first two days of Workshops and Special Interest Group Meetings were attended by ALTE Members and Institutional Affiliates, and the final day was an Open Conference Day for all those with an interest in language testing. The theme of the conference day was 'Developing and Implementing Language Tests for Younger Learners', chosen to reflect the increasing interest in encouraging young people to improve their competence in other languages.

The conference began with opening addresses from Johannes Ebert, the Secretary-General of Goethe-Institut, and Dr Michael Milanovic, ALTE Manager. These were followed by a presentation by Dr Shelagh Rixon, formerly of the University of Warwick, who drew on her research to show how course designers and test producers need to set reasonable age-appropriate goals for children and to ensure that there are well-planned yet changing approaches to learning as learners grow older.

Shelagh's presentation was followed by a presentation from Eli Moe from the University of Bergen who discussed the challenges of ensuring construct validity, reliability and positive washback when testing children. Henny Rønnepner from the Ministry of Schools and Further Education in North Rhine-Westphalia then spoke about how the Common European Framework of Reference for languages (CEFR) can provide orientation for foreign language learning in schools, and explained how in her state of North Rhine-Westphalia, all school leaving certificates show students' CEFR levels in modern foreign languages.

The afternoon sessions began with a presentation by Dr Neil Jones, Assistant Director, Research and Validation Group, Cambridge English Language Assessment and Director of the European Survey on Language Competences (SurveyLang), who discussed the findings of the survey and offered an interpretation in terms of implications for policy, suggesting how the need to improve language skills for employability in a globalised world may be reconciled with the promotion of linguistic diversity and intercultural dialogue.

Neil's presentation was followed by a series of short presentations by ALTE members and SurveyLang partners representing the Goethe-Institut, the University of Salamanca, the University for Foreigners, Perugia, and Centre international d'études pédagogiques (CIEP) who talked about some of the innovative approaches to language testing developed during the course of the project. The conference closed with a Round Table to discuss what could be learned from the SurveyLang project. This was moderated by Margot Kuzma from the European Commission, who was responsible for the project from the beginning.

Prior to the conference, ALTE ran a two-day course on **Assessing Young Learners** which was presented by Dr Szilvia Papp and Anthony King from Cambridge English Language Assessment, and also ran a one-day **Foundation Course in Language Testing: Getting Started**, which was run by Annie Broadhead, Consultant to Cambridge English Language Assessment.

## Forthcoming events

### ALTE 43rd Meeting and Conference, Salamanca, 17-19 April 2013

ALTE will hold its 43rd bi-annual Meeting and Conference at the University of Salamanca on 17-19 April 2013.

As with previous events, the first two days of meetings will be for representatives of ALTE members and institutional affiliates only, and the final day, Friday 19 April, will be an open conference day for all those with an interest in language testing. The theme of the conference day will be 'Language Assessment for Adults in the Context of Lifelong Learning' and speakers will include Dr Ardeshir Geranpayeh from Cambridge English Language Assessment, Clara de Vega and Gerardo Prieto from the University of Salamanca, and Professor Helen Spencer-Oatey from the University of Warwick.

Prior to the Meeting and Conference ALTE will run a two-day **Introductory Course on Action Research** on 15-16 April, and a one-day **Foundation Course on Language Testing: Getting Started** on 15 April.

For further information about all ALTE activities, please visit the ALTE website - [www.alte.org](http://www.alte.org). To become an Individual Affiliate of ALTE, please download an application form from the ALTE website or contact the Secretariat - [info@alte.org](mailto:info@alte.org). Individual affiliation to ALTE is free of charge and means you will receive advance information on ALTE events and activities and an invitation to join the ALTE electronic discussion forum.

## Announcement

It saddens us to inform our readers that Dr John Trim passed away on 19 January 2013. A pioneer in applied linguistics, John played a leading role in many important projects such as the Common European Framework of Reference (CEFR) and ALTE. John will be remembered for his significant contribution to applied linguistics and language education policy.

An interview with John Trim on the origin and evolution of the CEFR can be viewed at: [www.youtube.com/watch?v=BxGSQoxu6ho](http://www.youtube.com/watch?v=BxGSQoxu6ho)

**John Trim  
1924-2013**



This picture shows John with a special Cambridge Certificate which was presented to him as a tribute to his work in the University and in support of Cambridge English Language Assessment over a period of more than 50 years. This was presented to him at a dinner organised in his honour in October 2012.

Cambridge English Language Assessment will be publishing a tribute to John Trim in due course.



To subscribe to *Research Notes* and download previous issues, please visit:  
[www.cambridgeenglish.org/research-notes](http://www.cambridgeenglish.org/research-notes)

## Contents:

Editorial notes	1
Measured constructs: A history of Cambridge English language examinations 1913–2012 Cyril J Weir	2
Components of an elaborated approach to test validation Gad S Lim	11
A look at test taker characteristics for a C2 exam Ardeshir Geranpayeh	15
Consulting stakeholders as part of the Cambridge English: Proficiency exam revision Coreen Docherty and Debbie Howden	18
Development and construct of revised Cambridge English: Proficiency Ron Zeronis and Mark Elliott	22
A look into the future Michael Milanovic	31
2012 Research Notes reader survey Fiona Barker and Coreen Docherty	34
Caroline Clapham IELTS Masters Award 2012	37
Studies in Language Testing	38
ALTE report	39

For further information visit the website:  
[www.cambridgeenglish.org](http://www.cambridgeenglish.org)

Cambridge English  
Language Assessment  
1 Hills Road  
Cambridge  
CB1 2EU  
United Kingdom  
Tel. +44 1223 553997

