Examining Young Learners

Research and Practice in Assessing the English of School-age Learners
Examining Young Learners

Research and Practice in Assessing the English of School-age Learners

Szilvia Papp
Consultant, Cambridge Assessment English

and

Shelagh Rixon
Consultant, Cambridge Assessment English

With contributions from
John Field
CRELLA, University of Bedfordshire
Contents

Acknowledgements vii
Series Editors’ note ix
1 Introduction 1
   Szilvia Papp
2 The educational contexts of the teaching of English to young learners and the roles of assessment 18
   Shelagh Rixon and Szilvia Papp
3 Test taker characteristics of young learners 70
   Szilvia Papp
4 The cognitive validity of tests of listening and speaking designed for young learners 128
   John Field
5 The cognitive validity of reading and writing tests designed for young learners 201
   John Field
6 Introduction to context validity of tests for young learners 270
   Szilvia Papp
7 Context validity of listening tests of English for young learners 301
   Szilvia Papp
8 Context validity of speaking tests of English for young learners 341
   Szilvia Papp
9 Context validity of reading tests of English for young learners 382
   Szilvia Papp
10 Context validity of writing tests of English for young learners 411
   Szilvia Papp
Examining Young Learners

11 Scoring validity of tests of English for children and teenagers  
*Szilvia Papp*  
442

12 Criterion-related validity of tests of English for young learners  
Szilvia Papp  
510

13 Consequential validity of tests of English for young learners:  
The impact of assessment on young learners  
*Shelagh Rixon*  
547

14 Good practice guide  
*Shelagh Rixon*  
588

15 Conclusions and the way forward  
*Szilvia Papp*  
606

References 651
Appendix 1: Sample Cambridge English: Young Learners Listening tests 754  
Appendix 2: Sample Cambridge English: Young Learners Speaking tests 791  
Appendix 3: Sample Cambridge English: Young Learners Reading & Writing tests 806  
Appendix 4: Sample Cambridge English: Young Learners test from 2014 824  
Author index 827  
Subject index 835
Acknowledgements

This volume is about the examination system of one examination board, Cambridge Assessment English, but the authors’ hope and intention is that it should carry broader messages about the education of young people and their proper treatment, not least in the context of assessment. We have written the book in the conviction that a whole child approach is necessary, emphasising the well-being and social development of all school learners. In the modern age it seems that English language skills are one key to success but that linguistic prowess alone is not enough. English language skills need to be linked with 21st century and life skills including concept formation, ICT skills, critical reading and analysis, collaborative problem solving, creativity, self-regulation, emotional literacy, acting according to social norms, and learning to learn skills.

The volume has taken almost 10 years to write: it has gone through the initial teething problems, had several growth spurts, and gone through the intellectual revolution typical of school-age learners. We have learnt a lot about children and teenagers in the process.

The following people have played an important part in the development of the book:

• Mike Milanovic who, as Chief Executive Officer of Cambridge ESOL (as Cambridge Assessment English was then known) approved the original teaming of Szilvia and Shelagh as co-authors
• Evelina Galaczi and Ivana Vidaković as managing editors of the Studies in Language Testing series
• John Savage as Publications Assistant
• Cyril Weir as main reviewer
• many of the external and internal reviewers of the chapters who have made helpful suggestions that we were happy to take account of. Errors stemming from advice that we did not take or which we interpreted in our own way remain our responsibility.

Szilvia in particular, as original ‘in-house author’, would like to thank:

• Present and ex-colleagues at Cambridge English in the various teams (Research and Thought Leadership, Validation, Assessment) dealing with the Cambridge English: Young Learners tests and for Schools exams in various capacities: Neil Jones, Christine Walker, Mark Elliott,
Examining Young Learners

Michael Corrigan, Coreen Doherty, Lynne Stevenson, Glyn Hughes, Sarah McElwee, Guy Nicholson, Ed Hackett, Tony Green, Stuart Shaw, Dittany Rose, Jane Lloyd, Maggie Dunlop, and many others.

• Present and ex-colleagues outside Cambridge English too numerous to mention. Special mention is due to Sacha DeVelle, Iram Siraj, Sue Swaffield, Ayesha Ahmed, David Whitebread, Martin Johnson, and Clancy Blair, who have helped shape her ideas about assessing young learners.

• Her family and children without whose patience and support this volume would never have been finished and of whom she is so very proud.

Both authors would like to acknowledge the inspiration from colleagues and friends from the academic and language testing and assessment community who deal with young learners’ assessment: Marianne Nikolov, Annamaria Pinter, Angela Hasselgreen, Yuko Goto Butler, David Singleton, Tim Goodier, Veronica Benigno, Yeonsuk Cho, Aylin Ünaldi, Fiona Copland, Sue Garton, Wendy Arnold, Janet Enever, and many others.

We have sought permission for the use of the European Centre for Modern Languages’ 'Assessment of young learner literacy, linked to the Common European Framework of Reference for Languages' (AYLLit) descriptors in Table 11.8, and would like to thank Angela Hasselgreen for her kind permission to use the materials in Tables 12.1, 15.7 and 15.8.
Series Editors’ note

In the final chapter of Weir and Milanovic (Eds) (2003), the Studies in Language Testing (SiLT) volume on the *Cambridge English: Proficiency* (CPE) 2002 revision, the editors emphasised the need for Cambridge ESOL (now known as Cambridge Assessment English) to continue to keep abreast of and initiate research into the cognitive processes, strategies and performance conditions involved in the four macro-skills tested by its English language examinations. This was to be one of the core aims of the ‘constructs’ volumes in the SiLT series, the idea for which was already being discussed by Professor Cyril Weir and Dr Michael Milanovic with Lynda Taylor back in 2002.

*Examining Young Learners: Research and Practice in Assessing the English of School-age Learners* is the fifth volume in the Studies in Language Testing (SiLT) series to directly address the approach used by Cambridge English in its assessment of the four main language skills, the first being SiLT 26, *Examining Writing* by Shaw and Weir (2007), the second SiLT 29, *Examining Reading* by Khalifa and Weir (2009), the third, SiLT 30, *Examining Speaking* by Taylor (Ed) (2011), and SiLT 35, *Examining Listening* by Geranpayeh and Taylor (Eds) 2013. The first four volumes covered the examinations that formed the main suite of Cambridge English General English examinations. This volume completes the set of ‘construct’ volumes on the four macro-skills; together these five volumes constitute a significant endeavour in academic and publishing terms, representing more than 15 years of dedicated work among academics and practitioners working in the field of language test validation with and within Cambridge English.

The extensive authorial collaboration between external academics and Cambridge English staff which brought these volumes into being was facilitated by the strong endorsement and support firstly of Dr Michael Milanovic as Chief Executive of Cambridge ESOL, and more recently of Dr Nick Saville, the Director of the Research and Thought Leadership Group; this innovative initiative has resulted in a tangible and permanent resource which has immediate and lasting value for both the academic language testing community worldwide and the world of commercial language testing at the local, national, regional and international level.

The volumes are testimony to the academic rigour, experience and scholarship, as well as to the enormous expertise that resides within Cambridge English and among its external consultants, as well as to the organisation’s
Examining Young Learners

continuing commitment in the 21st century to generating validity evidence on how language constructs are measured by their language examinations.

They represent a significant contribution in the field to our theoretical understanding of the nature of the language abilities we are seeking to measure and to the practical approaches that test providers can adopt to achieve that goal with integrity, building a systematic, transparent and defensible body of validity argumentation in the process. Similar to efforts of the Educational Testing Service (ETS) in the late 1990s to develop the TOEFL 2000 Framework documents, the significant impact of these construct volumes on a wide variety of fronts should not be underestimated.

The perceived benefits of a clearly articulated theoretical and practical position for assessing listening/writing/reading/speaking skills in the context of the Cambridge English Qualifications are essentially twofold:

• Within Cambridge English: this articulated position deepens understanding of the current theoretical basis upon which Cambridge English assesses different levels of language proficiency across its range of products, and will inform current and future test development projects in the light of this analysis. It will thereby enhance the development of equivalent test forms and tasks.

• Beyond Cambridge English: it will communicate in the public domain the theoretical basis for the tests and hopefully provide a more transparent rationale for the way in which Cambridge English operationalises this in its tests. In addition, it provides a suitable framework for others interested in validating their own examinations, offering a principled basis and a practical methodology for comparing language examinations across the proficiency range. It therefore adds to the range of frameworks and models now available to test developers for analysing and describing the qualities of their tests and for guiding their research and validation activity.

Examining Young Learners: Research and Practice in Assessing the English of School-age Learners sets out to describe and evaluate how Cambridge English tests reading, writing, listening and speaking skills at different levels of ability in English as a second language, across the range of examinations it offers for young learners, covering the reference levels of the Common European Framework of Reference for languages (CEFR, Council of Europe) from pre-A1 to B2. As with the earlier Examining Writing, Examining Reading, Examining Speaking and Examining Listening volumes, this is achieved by presenting an explicit framework that structures the approach to validation according to a number of dimensions or parameters. This volume utilises the same theoretical framework which was originally proposed by Weir (2005a) and which seeks to take account of both the aspects of cognition, related to the mental processes the individual needs to engage in order to address a task,
and the features of language use in context that affect the ways in which a task is addressed. The authors also look at the practical assessment issues related to the marking and scoring of tests for young learners. Like its companion volumes therefore, this volume explores the triangular relationship between three critical internal dimensions of language testing tasks – the test takers’ cognitive abilities, the context in which the task is performed and the scoring process. Set alongside these are the twin external dimensions of consequential validity and criterion-related validity.

Cambridge English has been involved in the assessment of skills ever since it launched its first English language examination in 1913, the Certificate of Proficiency in English (CPE).

### 1913 examination

(i) Written:
(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

(ii) Oral:
Dictation ½ hour
Reading and Conversation ½ hour

A century ago, therefore, language learners wishing to certificate their command of English as a foreign or second language faced ‘an extremely demanding test of their abilities’ (Weir 2003:2), in which the testing of aural, oral, reading and writing language ability, both directly and indirectly, was integral to assessing their overall English language proficiency.

Since 1913, most new Cambridge English examinations have followed the model originally set by CPE and they have included an oral paper as an integral component of the language test battery, alongside tests of reading, writing, grammar and vocabulary; and later, in the 1970s, listening was to be tested in its own right. (For more details see accounts of other Cambridge tests in Hawkey 2004, 2009, and O’Sullivan 2006.)

The work of the Council of Europe in its Modern Languages programme, the emergence of the Threshold level, and the rise of the communicative language teaching movement all happened in the 1970s and 1980s and impacted on the Cambridge English approach to language testing. The 1975 revisions saw CPE taking a shape that in its broad outline is familiar to the candidate of today. The Listening and Speaking tests in particular represented major
developments on the 1966 revision and echoed the burgeoning interest in communicative language teaching in the 1970s. This era saw a change from teaching language as a system to teaching it as a means of communication as is detailed in Widdowson’s *Teaching Language as Communication* (1978).

In addition, an important study carried out in the late 1980s at Cambridge English was to have a powerful influence on the shape of things to come. Bachman, Davidson, Ryan and Choi (1995) carried out a Cambridge English-sponsored study entitled *An Investigation into the Comparability of Two Tests of English as a Foreign Language: The Cambridge–TOEFL Comparability Study* (published as SiLT 1). While ostensibly looking at the comparison between the First Certificate in English (FCE) and TOEFL in order to establish an empirical link between the level systems of each examination, this study actually ended up providing an in-depth critique of the Cambridge English approach with specific reference to the then well-developed and documented psychometrically oriented approach as instantiated by TOEFL. Significant issues in relation to reliability and validity emerged from the comparability study, which were addressed vigorously with the 1996 release of the revised FCE and subsequent release of the revised CPE in 2002.

As a result of Bachman et al’s study a much sharper focus on test construct definition and validation emerged. Where test construct had had to be pieced together post hoc from test specifications in earlier releases of FCE and CPE, there were now explicit statements on test construct. Measures were put in place not only to develop test content with systematic reference to the underlying construct but also to validate the nature of that construct. Additionally, the resulting work on standardising the test format was of major importance and instrumental in improving the reliability of the Cambridge English Qualifications.

More recently, we have also seen the emergence of the CEFR, which encourages examination providers to map their certification to the Framework. This volume examines how Cambridge English has approached this task in significant depth when exploring the criterion-related validity of its Young Learners tests. The approach taken by Cambridge English seeks not to establish the relationship with the Framework as a one-off study, but to deploy a methodology that ensures a long-term and continually verifiable relationship which is surely in the overall best interests of test users. Furthermore this volume evidences the efforts that are being made within Cambridge English to make the CEFR more relevant to young learners and so provide a better point of reference for criterion referencing and test development purposes.

Skills assessment at Cambridge English is now underpinned more formally than ever by a validation framework based on Weir (2005a); it builds on the work of Bachman (1989) which informed validation activities in the early 1990s, as well as on the Validity, Reliability, Impact, Practicality and Quality
(VRIPQ) approach developed by Cambridge English in the mid-1990s. The approach outlined in this volume not only allows Cambridge English to determine where current examinations are performing satisfactorily in relation to a range of relevant validity parameters; it also provides the basis for a programme of quality assurance and continuous improvement and for the construction of an ongoing research agenda. It provides an important benchmark against which test developers can evaluate the effectiveness of their respective approaches and it offers test users a model of what to expect from responsible examination providers.

The structure of the volume

Examining Young Learners: Research and Practice in Assessing the English of School-age Learners follows a similar structure to that which was successfully adopted for its sister publications Examining Writing (Shaw and Weir 2007), Examining Reading (Khalifa and Weir 2009), Examining Speaking (Taylor (Ed) 2011) and Examining Listening (Geranpayeh and Taylor (Eds) 2013). The outline shape closely follows their organisation with separate chapters on test taker characteristics; cognitive validity (two chapters on Listening/Speaking and Reading/Writing); context validity (five chapters, one generic and four specific on each of Listening/Speaking/Reading/Writing); scoring validity; consequential validity; and criterion-related validity. Apart from the two opening and the two closing chapters, each of the other chapters takes a component of the socio-cognitive validation framework to examine it in detail with reference to the Cambridge English: Young Learners tests, i.e. discussion of issues arising in the research literature on each component part is normally followed by consideration of Cambridge English practice in the area.

Examining Young Learners: Research and Practice in Assessing the English of School-age Learners is the product of collaboration between three acknowledged experts in language teaching, learning and assessment and in tests relating to young learners. They have been able to combine their extensive theoretical knowledge with practical application and expertise in these areas, and so provide valuable insights into the complex ecology that constitutes the assessment of young learners.

Cyril Weir
Nick Saville
Cambridge
June 2017