

| |
|---|
| <p>Institution: University of Bedfordshire</p> |
| <p>Unit of Assessment: 29 - English Language and Literature</p> |
| <p>Title of case study: Redefining English proficiency levels for second language education through applying our ground-breaking socio-cognitive framework for constructing and validating language tests</p> |
| <p>1. Summary of the impact</p> <p>Meaningful and useable definitions of language proficiency levels are essential for effective English curriculum design, language learning, teaching, and assessment. Since 2008 the socio-cognitive framework developed by the Centre for Research in English Language Learning and Assessment (CRELLA) has had a major impact on international test providers, enabling them to clarify the proficiency levels underpinning their English language tests, particularly the criterial features distinguishing one proficiency level from another. It has enabled them to develop more valid, dependable and fair measurement tools and to increase numbers of candidates taking their tests. For millions of successful candidates these enhanced English tests improve job prospects, increase transnational mobility and open doors to educational and training opportunities. Accurate proficiency tests lead to better informed and more equitable decision-making processes in society.</p> |
| <p>2. Underpinning research</p> <p>Building on the first comprehensive explication of the socio-cognitive framework (SCF) (publication Ref3.1 in 2007), CRELLA engaged in a long-term research programme to validate and refine the SCF. CRELLA staff systematically applied the SCF to internationally recognised English tests across different proficiency levels, skills and domains, to improve understanding of their construct validity. The goal was to investigate the extent to which these measurement instruments were an adequate and comprehensive representation of the real life construct – with special attention to the features that most efficiently distinguish one proficiency level from another. A full account of the research and its outcomes is reported in a coherent set of four academic volumes co-authored by CRELLA staff with colleagues in Cambridge English Language Assessment (CELA) (Refs3.1, 3.2, 3.3 and 3.5). This 8-year project provided for the first time a systematic and empirically based framework of criteria that can be used for analysing language tests and identifying areas where tests under-represent a construct or include test features that are irrelevant to it. The criteria are detailed and comprehensive, covering social, cognitive and scoring aspects of test design. The subsequent widespread application of the SCF to international high stakes tests has enabled the instruments to discriminate more effectively between different language proficiency levels (e.g., by the British Council, CELA, Trinity College London in the UK and the Language Training and Testing Centre (LTTC) in Taiwan).</p> <p>CRELLA's refined framework allows for principled, theoretical consideration of issues central to language test validity and for practical application in critical analyses of test content across the proficiency spectrum. It therefore has direct relevance and value to operational language testing/assessment contexts – especially where testing is conducted on an industrial scale. The SCF has largely superseded two earlier alternative frameworks. Bachman's 1990 Communicative Language Ability framework proved too difficult for examination boards to operationalise in large-scale practice. The Council of Europe's 2001 Common European Framework of Reference (CEFR) is insufficiently defined for test development purposes. Both lacked the essential cognitive dimension introduced by the SCF for discriminating between different levels of proficiency.</p> <p>From 2005-2007, Weir (Director of CRELLA 2005-present) mentored Shaw at CELA to apply the SCF initially to the exam board's writing proficiency tests (see academic volume on assessing writing – Ref3.1). This involved close linguistic analyses (lexico-grammatical, functional and discursal) of test input and output, from both sociolinguistic and psycholinguistic perspectives, in order to assess the tests' contextual and cognitive validity dimensions. This research enabled CRELLA to begin redefining the components of the SCF. From 2007 to 2009, Weir (at CRELLA) replicated the earlier approach with Khalifa (at CELA) using the exam board's reading proficiency tests (see 2009 volume on assessing reading – Ref3.2). CRELLA's ongoing research using the refined SCF generated two further academic research volumes (again with CELA staff) on assessing speaking and listening (Ref3.3 and Ref3.5). These drew on the additional research expertise and activities of four CRELLA members. Green (Principal Lecturer 2006-2008, Reader 2008-2013, Professor 2013-present) co-authored a chapter on <i>test taker characteristics</i> for the</p> |

Impact case study (REF3b)

speaking volume. **Field** (Senior Lecturer 2011-present) interpreted and operationalised *cognitive validity* from a psycholinguistic perspective for both volumes, based upon his ongoing research in cognitive processing. **Green's** research on the CELA speaking test rating scales informed the *scoring validity* component. Research into test washback and impact by **Green** and **Hawkey** (Senior Research Fellow 2006-2009, Visiting Professor 2009-present) provided the conceptual framework for the *consequential validity* component. **Taylor** (Senior Lecturer 2011-present) helped refine understanding of the value of the SCF for establishing second language proficiency levels through her contributions as author and editor to both the speaking and listening volumes.

A further major strand of research was conducted by CRELLA as a key founding partner of the *English Profile*, a long-term programme to provide the Council of Europe with Reference Level Descriptions that more fully define proficiency levels for English language education and assessment (see **Green's** academic monograph – **Ref3.4**).

3. References to the research

Ref3.1¹ Shaw, S.D. and **Weir, C.J.** (2007) *Examining Writing: Research and practice in assessing second language writing*. Studies in Language Testing 26, Cambridge: UCLES/Cambridge University Press. Reviewed in *Language Testing* 2010, 27: 141. 1,628 copies sold 2009-13

Ref3.2² Khalifa, H. and **Weir, C.J.** (2009) *Examining Reading: Research and practice in assessing second language reading*. Studies in Language Testing 29, Cambridge: UCLES/Cambridge University Press. Runner up in the 2012 Sage/ILTA triennial award for books on language testing. Reviewed in the *Modern Language Journal* 2011, 95/2: 334-335. 1105 copies sold 2009-13

Ref3.3² **Taylor, L.** (Ed.) (2011) *Examining Speaking: Research and practice in assessing second language speaking*. Studies in Language Testing 30, Cambridge: UCLES/Cambridge University Press. Includes individual chapters by **Taylor, Green, Field** and **Weir**. Reviewed on Teflnet, June 2012: edition.tefl.net/reviews/esl-exams/examining-speaking/. 3422 copies sold 2011-13

Ref3.4² **Green, A.** (2012) *Language Functions Revisited: Theoretical and empirical bases for language construct definition across the ability range*, Cambridge: UCLES/Cambridge University Press. 431 copies sold 2012-13

Ref3.5² Geranpayeh, A. and **Taylor, L.** (Eds.) (2013) *Examining Listening: Research and practice in assessing second language listening*. Studies in Language Testing 35, Cambridge: UCLES/Cambridge University Press. Includes individual chapters by **Taylor** and **Field**. 202 copies sold 2013

¹ Output submitted in RAE2008 exercise. CRELLA's RAE outputs were rated as 4* world-leading (20%), 3* internationally excellent (35%), 2* internationally recognised (25%), 1* nationally recognised (20%). ² Outputs submitted in REF2014 exercise.

Refs3.1-3.5 were commissioned by CELA from CRELLA staff, and four were co-authored with CELA colleagues. All were subject to external review by leading international experts in the field. Follow-on annual research funding (£96,275 from CELA since 2008 in support of research activities relating to the SCF and level definition) and numerous other joint publications with colleagues in CELA demonstrate clear pathways to impact. (See also *Research Notes: research.cambridgeesol.org/research-collaboration/research-notes* (Search construct definition).)

4. Details of the impact

Based upon the research outcomes reported in the set of 4 academic construct volumes (**Refs3.1, 3.2, 3.3, and 3.5** above), specific recommendations for improving Cambridge examinations were made in each publication's concluding chapter. CELA has been implementing the proposed changes in its examination revisions (2010-2013) to enhance their construct validity and thus their fitness for purpose. Examples of changes in the reading paper that have taken place include:

- SCF contextual and cognitive validity features for reading are now used to tag items in CELA's item bank to ensure tests constructed from one year to the next exhibit comparable validity indices, and guarantee that distinctions between proficiency levels are maintained.
- SCF analysis of tests led to removal of some tasks (gap filling) and inclusion of new tasks (intertextual summary) that are more cognitively valid for the highest proficiency level CPE.
- CRELLA's research on developing a methodology for establishing comparability in reading texts has now been incorporated into the organisation's guidelines for producing a test paper.

The *immediate* impact of CRELLA's ground-breaking research in test validation can thus be seen in improved measurement in high-stakes language proficiency tests (**Ref5.8**). By implementing the SCF for developing, validating, reviewing and revising its high-stakes examinations, CELA ensures the tests exhibit appropriate contextual and cognitive features at different language proficiency levels – see **Ref5.1**, an article by CELA staff member, Dr Gad Lim, who states: “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 result of this collaboration has been that CELA is able to define the constructs underlying Cambridge examinations at differing proficiency levels more explicitly. The SCF also ensures that due regard is paid to the psychological, physiological and experiential characteristics of the target test population so that tests are fair to all candidates irrespective of gender and background (see **Ref5.5**). In their 2009 article (**Ref5.3**), Khalifa and Ffrench (both directors at CELA) explain how the SCF is helping the organisation in its strategy to establish and confirm the alignment of its exams with levels of the CEFR for UKBA/QCA and other compliance purposes. CELA state that they promote CRELLA's four construct volumes among their stakeholders to support professional development and public accountability (**REF5.8**).

Since 2008, CRELLA's SCF has provided commercial test developers with the rigorous means of investigating and improving the overall fitness for purpose of their tests, thereby achieving wider impact in 2 key areas, which they attribute in part to the involvement and the research of CRELLA:

- **Contributing to economic prosperity**

Various test providers have enlisted CRELLA's expertise in applying the SCF to clarify the proficiency levels in their tests, improve their assessment products and thereby increase candidature and associated income. They include:

- Language Training and Testing Centre (LTTC) Taiwan – for the General English Proficiency Test (GEPT; 2 million additional candidates since 2008). **Ref5.2** refers to increased use of GEPT in Taiwan, e.g., 48% increase in Advanced stage 1 candidates between 2011 and 2013; the criterion related validation evidence generated by CRELLA; improvements to tests, specifications and distinctions between levels through applying SCF and **Weir** training its staff.
- Cambridge English Language Assessment (CELA) Examinations – for a suite of multiple, high stakes, general English language examinations across different levels and domains (rising from 2 million candidates in 2008 to over 4 million candidates in 2013, in 130 countries). See discussion above, client testimonial – **Ref5.8**, and website on underpinning research **Ref5.5**.
- British Council – for the new International Language Assessment (ILA) from 2010 (120,000 candidates pa) and the new Aptis tests from 2012 (150,000 candidates pa). See British Council report - **Ref5.6** and client testimonial **Ref5.9** for central role of the SCF in these.
- Trinity College London – for revision of Integrated Skills of English (ISE) tests, a suite of examinations at five levels, and validation of the Graded Examinations in Spoken English (GESE), a suite of examinations at twelve levels. Overall, Trinity tests are taken annually by 600,000 candidates worldwide. **REF5.10** for key role of CRELLA's SCF and staff in defining multiple levels of proficiency.

Valid, internationally recognised multi-level language examinations enable institutions and governments to handle recruitment, gate-keeping and transnational mobility in a well-informed way. For example, Cambridge English examinations are used by 13,000 employers, institutions and government ministries for admission, recruitment and training purposes, and immigration; they underpin critical selection and gate-keeping functions. CELA states that CRELLA's approach and contribution to their tests has been ‘particularly helpful in communicating the qualities of our tests for accreditation and recognition by agencies’ (Client testimonial **Ref 5.8**). Fit-for-purpose, well-regarded tests with a high value for successful candidates play an important part in improving revenue streams for their suppliers; a logical outcome in a commercial chain of events.

Accurate international language certification through accredited language tests plays an important role in economic prosperity. A recent British Council report (**Ref5.7**) evidences how, in the developing world, English language proficiency underpins the growth of national and individual wealth, and helps drive economic development. Reducing unemployment is seen as a means of

Impact case study (REF3b)

securing political stability. Governments in the developing world view the improvement and certification of English language skills at all levels as an essential part of achieving growth, by giving domestic companies a competitive edge in the global economy. They are also a significant factor in attracting Foreign Direct Investment. Coleman (**Ref5.11**) provides empirical evidence of English: increasing individuals' employability; enabling international collaboration and cooperation; providing access to research and information; facilitating the international mobility of students, tourists, workers, etc. He concludes that English impacts on individuals, on particular industrial sectors (especially service economies) and at a national level. Valid certification of English proficiency through language tests enhanced by CRELLA's SCF plays an important role in this.

- **Enhancing educational and employment opportunities**

Cambridge English examinations, enhanced by the application of CRELLA's SCF, are used by employers around the world as evidence of candidates possessing the requisite levels of proficiency in English. For example, more than 3,000 educational institutions, businesses, government departments and other organisations around the world now recognise the Cambridge Certificate in Advanced English (Cambridge Advanced) as a quality index of high-level achievement (Cambridge website **Ref5.5** and client testimonial **Ref5.8**). Improved Cambridge examinations have had a direct impact on teaching and a positive impact on education. External certification through CELA examinations has contributed to more efficient English language learning in primary and secondary schools with faster progression up the CEFR scales of ability. In the Italian State System, for example, empirical evidence shows a rise of a full CEFR level among secondary school leavers (from Cambridge English Preliminary (PET) to Cambridge English First (FCE)); increased student motivation for learning English; greater parental satisfaction; and improved pedagogical practice (see research study **Ref5.4** and client testimonial **Ref5.8**).

Research into proficiency levels informed CRELLA's development of the International Language Assessment (ILA), a British Council web-based test used to place c.120,000 learners of all proficiency levels into appropriate classes in its teaching centres worldwide more efficiently and effectively with fewer false placements than before (see client testimonial **Ref5.9**). In 2012 CRELLA's SCF provided the conceptual basis for the development of the new British Council Aptis testing service used by corporate businesses, government organisations, educational institutions and NGOs worldwide for: benchmarking students and employees; language audits to identify training needs; filtering current/potential employees for promotion/interview, and as a diagnostic tool to identify strengths/weaknesses of people seeking employment (see research report **Ref5.6** and client testimonial **Ref5.9**). Reem Salah, GlaxoSmithKline, Egypt, describes how: "Aptis has allowed us to benchmark our employees' English skills easily and affordably. We have been able to identify those employees who need further training, and those who may be suitable for an alternative role within our business."

5. Sources to corroborate the impact

Ref5.1 Lim, G.S. (2013) Components of an elaborated approach to test validation, *Research Notes*, 51, February, pp.11-14: www.cambridgeenglish.org/Images/130828-research-notes-51-document.pdf

Ref5.2 Director, Language Training and Testing Centre (LTTC) Taiwan, GEPT testimonial

Ref5.3 Khalifa, H. and French, A. (2009) Aligning Cambridge ESOL examinations to the CEFR: issues and practice. *Cambridge ESOL Research Notes*, 37, August, pp.10-14: www.cambridgeenglish.org/Images/23156-research-notes-37.pdf

Ref5.4 Hawkey, R. et al. (2013) *The Progetto Lingue 2000 Revisited (PLISR), an impact study* Research report for Cambridge English Language Assessment

Ref5.5 www.cambridgeenglish.org/research-and-validation/fitness-for-purpose/(search socio-cognitive)

Ref5.6 www.britishcouncil.org/sites/default/files/documents/aptis-test-development-approach-aug-2012-1.pdf

Ref5.7 www.britishcouncil.org/new/Documents/full_mena_english_report.pdf

Ref5.8 Director, Cambridge English Language Assessment Examinations testimonial

Ref5.9 Senior Adviser British Council for Aptis and ILA testimonial

Ref5.10 Head of Academic Research, Trinity College London Examinations testimonial

Ref5.11 Coleman, H. (2010) *The English Language in Development*. British Council: London