

<b>Institution:</b> University of Bedfordshire		
<b>Unit of Assessment:</b> 27 English Language and Literature		
<b>Title of case study:</b> Refining English language proficiency assessment for the digital age: applying a comprehensive framework		
<b>Period when the underpinning research was undertaken:</b> 2005-2020		
<b>Details of staff conducting the underpinning research from the submitting unit:</b>		
<b>Name(s):</b>	<b>Role(s) (e.g., job title):</b>	<b>Period(s) employed by submitting HEI:</b>
1. Sathena Chan	1. Senior Lecturer	1. 2013-present
2. John Field	2. Reader	2. 2011-2020
3. Anthony Green	3. Professor	3. 2006-present
4. Roger Hawkey	4. Senior Research Fellow	4. 2007-2016
5. Nahal Khabbzbashi	5. Senior Lecturer	5. 2017-present
6. Fumiyo Nakatsuhara	6. Reader	6. 2009-present
7. Lynda Taylor	7. Senior Lecturer	7. 2011-2015
8. Cyril Weir	8. Powdrill Chair Professor	8. 2005-2018
<b>Period when the claimed impact occurred:</b> 1 August 2013 - 31 July 2020		
<b>Is this case study continued from a case study submitted in 2014?</b> N		
<b>1. Summary of the impact</b>		
<p>Accurate educational measurement leads to effective English curriculum design, language learning, teaching, assessment, and better informed, more equitable decision-making in society. For millions of language learners, enhanced English assessments, based on CRELLA's research, have improved job prospects, increased transnational mobility, and opened doors to educational and training opportunities. CRELLA's socio-cognitive framework (SCF) is applied to a new digital generation of assessments, meeting local needs both geographically and professionally. It informs training for teachers and assessment professionals; supports equity and diversity, with more learners worldwide able to access quality English language assessments; and increases the value of assessments as a basis for learning.</p>		
<b>2. Underpinning research</b>		
<p>Responding to the urgent need for a practical, theoretically-sound framework for language test development and validation, research undertaken by the University of Bedfordshire's Centre for Research in English Language Learning and Assessment (CRELLA) has resulted in the conception and development of the socio-cognitive framework (SCF) (3.1). SCF was the culmination of a thorough review of concepts of test validity by <b>Weir</b>, CRELLA's founding director and Powdrill Chair Professor of Applied Linguistics. It was based on his 30-year engagement in research in more than 30 countries. SCF allows for principled, theoretical consideration of issues central to language test validity and for practical application in critical analyses of test content. It therefore has direct relevance and value to operational language testing/assessment contexts – especially where testing is conducted on an industrial scale. Since its inception, CRELLA has continued to develop and refine SCF through its application to major testing programmes. The Centre's research goal has been to investigate how assessments can provide comprehensive representation of real-life language use – with special attention to the features that most efficiently distinguish one proficiency level from another. Most recently, this has involved adapting and expanding SCF for use in novel assessments that are increasingly delivered in digital environments. For example, digital test delivery makes it practicable to divide, pause and play back recordings of online speaking tests. In response to</p>		

this, **Khabbazzbashi** (Senior Lecturer specialising in Digital Speaking Assessment) introduced a novel dimension of ‘scoring performance by test part’ to the SCF category of scoring mode (3.5). CRELLA research has demonstrated the operationalisation of SCF through work with the UK’s leading English language testing organisations. This has informed test design and scoring systems that meet current local and global language learning needs (3.2, 3.3). Authored by CRELLA staff in collaboration with colleagues in Cambridge Assessment English (CAE), four key volumes reported the outcomes of early SCF research. These volumes covered the application of SCF to various proficiency levels in Writing, Reading, Listening, and Speaking across CAE examinations. They investigated how test tasks posed at different levels reflected the linguistic, interactional and cognitive demands placed on language learners in real-life contexts (3.2). Important components of this include chapters part-authored by CRELLA staff on their research into different components of the SCF: **Green** (current Director of CRELLA and Professor in Language Assessment) on test-taker characteristics, **Hawkey** (a former Senior Research Fellow) on consequential validity, and **Field** (Reader in Cognition in Language Learning) on the cognitive validity of CAE examinations.

Research by **Weir** and O’Sullivan, Head of Assessment Research & Development at the British Council (BC), investigated the contribution of BC to language testing theory and practice (3.3). It explains how SCF guided BC in developing a unique digital and localised approach to language assessment which captures differences in language use for specific social purposes and in local cultural contexts. One notable development is BC’s computer based *Aptis* testing system, launched in 2012, and its distinctive variants targeting different candidate groups, such as *Aptis for Teachers* (launched in 2013) and *Aptis for Teens* (launched in 2015).

CAE and CRELLA are partners in the English Profile global research programme funded by Cambridge Assessment English and Cambridge University Press and, initially, through the Lifelong Learning Programme of the European Commission (505491-LLP-1-2009-1-UK-KA2-KA2NW: 2009-12 EUR448,730 – c.GBP390,000). Within this, **Green** (2012), **Nakatsuhara** (2018; Reader, specialising in Spoken Discourse), and **Chan** (2018; Senior Lecturer, specialising in Integrated Assessment of Reading and Writing) have published research monographs in the English Profile Research Series, all of which have defined more precisely levels of proficiency in English in relation to the Council of Europe’s (2001) Common European Framework of Reference for Languages (CEFR). This collaborative work has been extended through the investigation of a new generation of assessments encompassing digitally mediated tasks. An example is **Khabbazzbashi** and Galaczi’s (2020) study (3.5) which compared different models for scoring spoken performance, with respect to scoring validity, a crucial component of SCF. This informed the automated marking of digitally delivered speaking tests.

**Field** (3.4) added to the cognitive profiling within SCF by incorporating evidence on the stages of early cognitive development into the criteria for grading the language skills of young learners. The SCF parameters expanded for young learners by Field (3.4) have also been applied to a major international test revision and development project for primary and secondary students in Uruguay. CRELLA was granted GBP91,367 for the revision of the national computer-adaptive English test (covering vocabulary, grammar, reading and listening) and the development of a new online speaking test for young learners of English.

SCF was also extended to inter-test comparison research in the professional domain. **Taylor** (Visiting Professor, specialising in the Impact of Testing on Society) and **Chan** (3.6) identified and refined SCF parameters to scrutinise a range of standardised English proficiency tests available for use within registration procedures for overseas-trained doctors seeking to practice in the UK. The CRELLA research team was funded by the UK General Medical Council (Project ID: GMC133: GBP82,685) to help them evaluate the tests’ suitability for their own professional registration purposes and make decisions on acceptable tests and scores.

### 3. References to the research (indicative maximum of six references)

**3.1 Weir, C. J. (2005). *Language testing and validation*. Basingstoke: Palgrave Macmillan.**  
[This book presents the original socio-cognitive framework (SCF), detailing each component and relevant parameters while referring extensively to the literature in the fields of language testing and educational measurement.]

**3.2** Geranpayeh, A., & Taylor, L. (Eds.). (2013). *Examining listening: Research and practice in assessing second language listening*. Cambridge: Cambridge University Press.

[One of four books in the Studies in Language Testing Series co-authored by CRELLA members, which exemplifies how SCF is embedded in every stage of test development and validation by Cambridge Assessment English. Chapters in this book were authored by **Green, Hawkey, Field, Taylor and Weir.**]

**3.3** Weir, C. J., & O'Sullivan, B. (2017). *Assessing English on the global stage: The British Council and English language testing, 1941-2016*. Sheffield: Equinox.

[Investigates the unique contribution of BC to language testing theory and practice. Gives full recognition to the informing principles of the SCF. Chapter 6 (pp.257-330) reflects fundamental role of SCF in test development and validation at BC, especially that of their digitally delivered test, Aptis.]

**3.4** Field, J. (2018). The cognitive validity of tests of listening and speaking designed for young learners. In S. Papp & S. Rixon (Eds). *Examining young learners: Research and practice in assessing the English of school-age learners*. Studies in Language Testing vol. 47 (pp. 128-200). Cambridge: Cambridge University Press.

[Considers the role of cognitive development when designing listening and speaking tests for young learners. This research fed into the development of the digitally delivered national speaking test for Uruguayan primary and secondary school children (see Section 4).]

**3.5** Khabbazzashi, N. & Galaczi, E. D. (2020). A comparison of holistic, analytic, and part marking models in speaking assessment. *Language Testing*, 37(3), 333-360.

DOI:10.1177/0265532219898635

[Compares marking models in an online speaking test. Led to refinement of SCF scoring parameter]

**3.6** Chan, S. & Taylor, L. (2020). Comparing writing proficiency assessments used in professional medical registration: a methodology to inform policy and practice. *Assessing Writing*, 46. DOI:10.1016/j.asw.2020.100493

[Based on a project commissioned in 2014-15 by the UK General Medical Council (GMC) to help them evaluate the suitability of various English language proficiency tests available for use in licensing overseas trained doctors to operate in the UK.]

#### 4. Details of the impact

Accurate educational measurement leads to effective English curriculum design, language learning, teaching, assessment, and better informed, more equitable decision-making in society. The immediate impact of CRELLA's cutting-edge research in test validation can be seen in improved measurement of proficiency in several high-stakes tests of English language proficiency. The socio-cognitive framework (SCF) has enabled numerous English language test providers around the world to improve test quality and to clarify the proficiency levels underpinning their tests. The effect has been to enhance their reliability and reputation, leading to increased transparency for employers who use the tests. Through validating digital language assessment, CRELLA's work has provided prospective students and jobseekers a wider and more affordable range of English language tests from which to choose, beyond traditional face-to-face, in-person testing. In short, CRELLA's work on assessment has opened more doors thus affording greater educational and employment opportunities to successful candidates. Below we present the examples of two of the UK's leading providers of English language assessments, Cambridge Assessment English (CAE) and the British Council (BC): both base their test design principles on SCF.

The Assistant Director (Assessment) at CAE observes that, "Over the past 20 years, Cambridge English has established a close working relationship with CRELLA and has embedded the socio-cognitive framework (SCF) for language test development and validation in its operation" (5.1). For example, SCF's contextual and cognitive validity features for reading are now used to tag items in CAE's database of test material, ensuring that individual test papers consistently and

accurately represent distinctions between proficiency levels and present a similar experience across for users. The impact of the accurate measurement of English proficiency is substantial: “Cambridge Assessment English qualifications are taken by more than 7 million people every year in 130 countries and recognised by more than 25,000 universities, employers and governments” (5.1).

SCF has also provided the theoretical foundations for new digital CAE tests launched in the current REF period, such as Linguaskill, released in 2017. “As of December 2020, this new digital test has been administered over 375,000 times and is recognised by 450 organisations across 52 countries” (5.1). The practical value of SCF for assessment design and validation has led to its use in other specialised fields; Cambridge Assessment Admissions Testing employs SCF for the biomedical admissions test (BMAT) (5.2), recognised by 40 medical universities worldwide.

Like CAE, BC has implemented SCF in developing, validating, reviewing, and revising its English language tests. Since 2012, SCF has provided the conceptual basis for the digitally based Aptis testing service, used by businesses, governments, educational institutions and NGOs in more than 85 countries, and is used for gatekeeping and diagnostic purposes. According to the Senior Researcher for the BC’s Assessment Research Group, “In the 2019-2020 financial year, approximately 300,000 tests were delivered in 100 countries” (5.3). He notes that SCF has helped BC to ensure that the tests exhibit appropriate contextual and cognitive features at different language proficiency levels through:

- test specifications that guide item writers in targeting their questions to learners at different proficiency levels
- rating scales used in awarding scores on the writing and speaking components
- examiner training procedures designed to ensure accurate and consistent judgement
- a research programme designed to ensure that the tests are and remain fit for purpose.

The British Council’s internal continuing professional development programme is based on SCF (5.3, 5.4) and SCF provides the theoretical grounding for its ‘How Language Assessment Works’ initiative, which includes a free online FutureLearn course on assessment issues for a global audience. This has been taken by 45,315 people in more than 150 countries to date (5.3).

In addition to the CAE and BC testing programmes, CRELLA has used SCF in conducting, supporting and advising on a range of test development initiatives, for new markets in the UK and around the world. Examples include:

- EnglishScore: an innovative English language test designed by CRELLA and now available in 150 countries, administered for free via mobile devices and scored in under 40 minutes. According to the BC EnglishScore Ltd CEO, (5.5), since launching in 2019, over 1.7 million users have downloaded the EnglishScore app from the GooglePlay store, approximately 80,000 tests are taken each month and the organisation now employs 19 members of staff (headcount: 19, FTEs: 18). He adds that, “Without the input and expertise of CRELLA there is no way that we could have made the progress that we have.”
- The National English Adaptive Test (NEAT) and the Plan Ceibal Online Speaking Test in Uruguay: Commissioned by the Uruguay government in 2018, CRELLA re-developed NEAT and designed a new online speaking test to better measure the progress of primary and secondary school children. Since 2019, NEAT has been administered to over 75,000 children and the Plan Ceibal Online Speaking Test to over 6,000 children nationwide. According to the programme’s director, “The application of the SCF informed quality assurance procedures for the design, piloting, validation and large-scale digital administration of the tests. It helped us to ensure that test tasks were suitable for online delivery, appropriate for young learners in Uruguay, and that the rating scales used to judge performance reflected the local context.” (5.6).

Furthermore, using the components and parameters of the SCF, CRELLA devised an effective methodology for comparing different English language tests. This research was commissioned by the General Medical Council (GMC) to help them evaluate the suitability of various English language tests available for use in licensing overseas trained doctors to operate in the UK. As a direct result of CRELLA’s research, the GMC now recognises another test alongside IELTS: the Occupational English Test (available at more than 145 test venues in 44 countries) (5.7).

The wider impact of the SCF is evidenced by various educational outreach initiatives taken by CRELLA staff, supporting quality assessments for new markets, and developing the assessment



skills of teachers and assessment professionals in the UK and globally. **Weir, Green, Taylor, Field, Khabbzbashi** and **Nakatsuhara** have offered SCF-informed assessment training courses and webinars to 24 international institutions. These include the Association of Language Testers in Europe (ALTE: an association of test providers for all major European languages including the Goethe-Institut, Alliance Francaise and Instituto Cervantes) (**5.8**), the International Language Testing Association (ILTA), and the International Association of Teachers of English as a Foreign Language (IATEFL).

Using SCF as the informing theory, two large-scale EU-funded projects, in which CRELLA was a partner, have enhanced the assessment literacy of English teachers in Eastern Europe and the Russian Federation. With partners across Europe, the Erasmus+ TALE project (2015-1-CY01-KA201-011863: EUR268,884, c.GBP235,000) produced an on-demand, open-access course for language teachers designed to help them make effective use of assessment in the classroom.

**Green** and **Hamp-Lyons** authored three of the eight course modules. The project was awarded the BC Innovation in Assessment Prize in 2019. In a series of workshops in Ukraine conducted in association with Taras Shevchenko University, teacher trainers were trained by CRELLA staff in testing principles and new university tests in English were developed.

Similarly, in an EU (Tempus) funded project that ran from 2011 to 2014 (517114-TEMPUS-1-2011-1-UK-TEMPUS-SMHES, EUR910.860, c.GBP790,000), led by **Green**, eleven Russian universities sent teacher trainers to CRELLA to work together in designing a new course in language assessment for preservice and in-service teachers. The handbook for the course was published by Cambridge University Press ([https://www.beds.ac.uk/media/kmrjwwzw/proset\\_lta-course-brochure\\_en.pdf](https://www.beds.ac.uk/media/kmrjwwzw/proset_lta-course-brochure_en.pdf)). Participants also upgraded their test writing and development skills to improve the quality of the Russian national state examination in English. The Dean of the Faculty of Foreign Languages at Mordovia State University and the coordinator for the Russian partners in the project (**5.9**) testifies that the SCF and CRELLA's research "has had a profound and sustained impact on language education in the Russian Federation." The training courses are still run regularly by all the project partners: "In the 2014-2020 period, over 20,000 teachers have received in-service assessment training through programmes developed by the project."

## 5. Sources to corroborate the impact (indicative maximum of 10 references)

**5.1.** Testimonial from Assistant Director (Assessment), Cambridge Assessment: English. *Provided as PDF.*

**5.2.** Cheung, K. Y. F., McElwee, S., & Emery, J. (2017). *Applying the socio-cognitive framework to the Bio-Medical Admissions Test (BMAT)*. Cambridge: Cambridge University Press. [SCF is used throughout the book, but Chapter 1 (pp.1-16) first mentions and discusses it]. *Provided as PDF.*

**5.3.** Testimonial from Assessment Manager, British Council Assessment Research Group. *Provided as PDF.*

**5.4.** O'Sullivan, B. (2015). *Technical report: Aptis test development approach*. *Provided as PDF.* [https://www.britishcouncil.org/sites/default/files/tech\\_001\\_barry\\_osullivan\\_aptis\\_test\\_-\\_v5\\_0.pdf](https://www.britishcouncil.org/sites/default/files/tech_001_barry_osullivan_aptis_test_-_v5_0.pdf)

**5.5.** Testimonial from Chief Executive, BC EnglishScore Ltd. *Provided as PDF.*

**5.6.** Testimonial from Manager Plan Ceibal, and Director of Ceibal en Ingles, Uruguay. *Provided as PDF.*

**5.7.** General Medical Council (2018). *GMC to accept new English language qualification for non-UK doctors*. <https://www.gmc-uk.org/news/news-archive/gmc-to-accept-new-english-language-qualification-for-non-uk-doctors> *Provided as PDF.*

**5.8.** Testimonial from ALTE Secretary General. *Provided as PDF.*

**5.9.** Testimonial from Dean of the Faculty of Foreign Languages, National Research Mordovia State University. *Provided as PDF.*