



BRITISH COUNCIL PERSPECTIVES ON ENGLISH LANGUAGE POLICY AND EDUCATION

Language Testing as Cultural Relations

Barry O'Sullivan, Mina Patel and Gary Mundy

British Council Perspectives on English Language Policy and Education

The Perspectives Series sets out the British Council's approach to a range of issues around the English language in learning systems across the world. The papers build on the experience of British Council experts throughout the organisation. Contributions come from the English Language Research Group (ELRG) which includes representatives from the Assessment Research Group, English Teaching and Learning, as well as English in Education Systems (EES).

© British Council 2021.
ISSN 2752-5562

Web: www.britishcouncil.org/exam/aptis/research/publications/british-council-perspectives-english-language-policy-and-education

Contents

Authors	3
1. The British Council and Cultural Relations	4
2. A Short Historical Contextualisation	5
3. Language Testing and Assessment.....	7
4. The British Council and Language Assessment	8
5. Linking the British Council's Cultural Relations Mission with its Assessment Activities	11
5.1 The China Case Study.....	12
5.2 The New Directions Case Study.....	13
5.3 The English Impact Case Study	15
5.4 The Singapore Case Study.....	17
5.5 Commentary	19
6. Building Our Evidence: An Integrated Approach	19
6.1 Our Cultural Relations Evidence Needs	20
6.2 The Types of Evidence We Need for Cultural Relations	20
6.3 The Need for an Integrated Approach to Cultural Relations Evidence	22
7. Conclusion.....	25
References	27

Authors

Barry O'Sullivan

Professor Barry O'Sullivan OBE is the British Council's Head of Assessment Research & Development. He has worked on numerous test development and validation projects globally and advises ministries and institutions on assessment policy and practice. Barry has undertaken research across many areas on language testing and assessment, and has worked on the development and refinement of the socio-cognitive model of test development and validation since 2000. He regularly speaks at international conferences and has over 100 publications. He is the founding president of the UK Association of Language Testing and Assessment, and is an elected fellow of the Academy of Social Sciences and of the Asian Association for Language Assessment.



Mina Patel

Mina Patel is Assessment Research Manager with the Assessment Research Group at the British Council. Her background is in English language teaching and training. She has worked in the UK, Greece, Thailand, Sri Lanka and Malaysia as a teacher, trainer, materials developer and ELT projects manager, and has extensive experience working with ministries of education in East Asia. Mina has presented at numerous national and international conferences on ELT related matters. Mina's interests in language testing and assessment lie in the areas of language assessment literacy and the impact of testing and assessment. She is currently a PhD student with CRELLA at the University of Bedfordshire, UK.



Gary Mundy

Gary Mundy is the British Council's Senior Regional Evaluation Advisor in East Asia. He has almost twenty years of operational experience in the production and use of evidence to understand and drive greater impact for programs in global health, education, and governance. Gary has worked extensively in Asia and sub-Saharan Africa, implementing a wide range of methodological approaches to understand what works, for who, where and under what circumstances. He has presented and published work at a wide range of conferences and in research/evaluation journals.



1. The British Council and Cultural Relations

The mission of the Royal Charter, which lies at the core of everything the British Council does, states that the formal role of the organisation is in “promoting a wider knowledge of [the UK] and the English language abroad and developing closer cultural relations between [the UK] and other countries”. Just how the organisation views cultural relations is clear from the following definition from the Goethe-Institute and the British Council (2018, p. 7).

Cultural relations are understood as reciprocal transnational interactions between two or more cultures, encompassing a range of activities conducted by state and/or non-state actors within the space of culture and civil society. The overall outcomes of cultural relations are greater connectivity, better mutual understanding, more and deeper relationships, mutually beneficial transactions and enhanced sustainable dialogue between people and cultures, shaped through engagement and attraction rather than coercion.

How this definition is operationalised in the real world depends on the needs and expectations of the stakeholders involved. In his overview of cultural relations at the British Council, Singh (2020, p. 7) argues that the:

British Council defines its role as supporting the prosperity, security, and influence of the United Kingdom through its cultural relations activities. The notion of cultural relations is expansive, including eight important ways in which cultures interact through arts, English, education and science, skills and enterprise, young people, civil society and justice, testing and assessment, and women and girls.

It is clear from the above that all British Council activities, including its work in testing and assessment, can be seen to contribute in some way to the organisation’s cultural relations agenda. In this paper, we set out how British Council’s work in testing and assessment has operationalised the cultural relations construct as highlighted in the above quotations. We support this with evidence from the organisation’s activities in the area in recent decades.

2. A Short Historical Contextualisation

The earliest example of a formal competitive examination is the Kējǔ, the Chinese Imperial Examination system, which had its origins in the Han Dynasty (202 BCE–220 CE), was standardised by the Sui (581–618) and ended in 1905 during the last dynasty, the Qing (1644 CE–1912 CE). The ultimate purpose of the Kējǔ, which consisted of increasingly challenging tests at the local level, regional, national and, ultimately, palace level, was to identify talented individuals to act as Mandarin, or senior civil servants, to manage the empire.

By controlling all aspects of the management of the Empire, the Sui, and later dynastic regimes, set out to ensure that even those at the very edges of their empire were fully acculturated, as can be seen in the following quotation from 1425, as translated by Wade (2005).

If prefectural, sub-prefectural and county instructors are appointed to guide the training of the native people, to provide correct models and to demonstrate the ways of education, then day-by-day, the students will be thoroughly imbued with what they see and hear. Thereby, their 'man' and 'yi' [i.e. foreign or barbarian] ways will be changed and again they will take on the ways of the Chinese. 'This would certainly be a wonderful thing!' the Emperor said: The ancients said: 'There should be no differentiation in employing the worthy'. They also noted: 'In teaching there should not be any distinction between classes'. Teachers act as models. Using Chinese ways to change barbarian ways – there is nothing more important than this. Order the Ministry of Personnel to select suitable personnel and make the appointments.

This 'civilising' practice spread as the influence of the Chinese Empire grew to encompass to what Fairbanks (1968, p. 2) calls the Sinic Zone, comprised of mainly nearby territories who shared a similar culture: Korea, Vietnam, the Ryūkyū Islands (part of Japan since 1879), and for a relatively short time, Japan itself. Interestingly, the system was long-lasting in Korea (958 CE–1894 CE) and Vietnam (1075 CE–1913 CE), although it never fully took root in Japan, most likely due to long-established local practices of promotion systems that were based on the status and social standing of a person's family.

Following the foundation of the Peoples' Republic of China in 1949, the study of the Kējǔ fell out of favour. However, in recent years there has been a noticeable change in the position of the Chinese government. The fact that the Kējǔ is recognised globally as the precursor of all modern testing systems has clearly prompted them to recognise the cultural capital it brings to China and the cultural relations benefits that come with it. So, once proudly exported to tributary states as a symbol of its advanced civilisation, the Kējǔ has now come to represent the fifth of China's significant contributions to human civilisation (the other four being papermaking, the compass, gunpowder, and printing).

'By highlighting the positive potential of the test in opening doors for the individual test-taker, the idea of mutual benefit becomes clearer to all stakeholders. It is through this concept of mutual benefit that cultural relations at the level of the individual participant enters the equation.'



3. Language Testing and Assessment

The terms ‘testing and assessment’ are regularly used interchangeably in the education literature. However, for the purpose of this paper, we will use the two in quite different ways. Here, ‘assessment’ refers to the non-judgemental evaluation of one or more aspects of an individual learner’s language proficiency for primarily developmental purposes. ‘Tests’, on the other hand, are judgemental in nature and can be low or high stakes. Low-stakes tests, e.g., for placement on a course of study, are often free or very affordable for the user, with no major life-changing decisions based on their outcomes. High-stakes tests, however, are typically costly and can have major consequences for all involved.

While we will focus primarily on high-stakes tests, we recognise that lower-stakes tests and assessments can also offer significant cultural benefits by highlighting the expertise of an institution or industry associated with a particular nation or the openness of these to share the underlying knowledge with other institutions of governments.

A high-stakes test is usually seen as an isolated, transactional activity. A test-taker pays for a test, takes it, receives the result and that is the end of the transaction. However, to see a test like this is to diminish its value and contribution to a variety of impact agendas, including that of cultural relations. The way the test is constructed, the types of questions it contains, the choice of language within those questions, and the marking system all represent the educational philosophy of an organisation or a context in which the test has been developed, reflecting the standards and values that are embedded in the culture of the developing organisation and/or country. In addition, a test, though important in itself, is one part of a much bigger educational system or process that exists to support individuals, institutions and countries achieve their goals and ambitions.

Seen through the eyes of an individual test-taker, the process begins when deciding which test to take. Factors that impact this include the reputation of the test, recognition by receiving organisations, customer service, the mode and standards of delivery and, of course, the cost. Test-takers do not necessarily want to take a test. They want the test to allow them to attain cultural (social or educational) and/or financial benefit. Test developers should therefore learn to recognise that their product is offering the test-taker a means to an end and make this clear in all messaging to prospective test-takers. Of course, this proposal comes with a caveat: if the developer is not seen to support the test-taker in his/her journey then there is a risk of rejection, and in these days of social media, such attitudes can spread rapidly and significantly impact on the number of test-takers willing to take the test (even if the information upon which the attitude is based is false). The image portrayed in this transaction is typically that the test is a hurdle in the test-taker’s path. By changing the narrative, to highlighting the positive potential of the test in opening doors for the individual test-taker, the idea of mutual benefit becomes clearer to all stakeholders. It is through this concept of mutual benefit that cultural relations at the level of the individual participant enters the equation.

At an institutional level, the cultural relations aspect of assessment works on a number of levels. Where the role of the institution is solely that of a Receiving Organisation (RO) which uses the scores from a test, there is an underlying acknowledgement of the expertise of the developing organisation combined with the confidence that the test can be fully trusted to allow the RO to make fair and meaningful decisions (e.g. in terms of selection or promotion). In this way, the test can be seen as a facilitator for a cultural relations outcome, with both parties benefiting from the use of the test.

Where there is a broader connection between the test developer and an institution, the possibility of a positive cultural relations outcome increases significantly. A good example of this is where the developer works directly with the institution to grow its capacity to build tests of a high quality. The understanding and collaboration shared between the protagonists benefit both parties as both bring particular expertise to the interaction – the developer in the area of testing, the institution in the knowledge and understanding of the context of use and its key stakeholders.

On a national level, cultural relations are most likely to prosper in response to long-term, mutually beneficial relationships and trust building. The cases cited here are examples of how the British Council, working directly with a local ministry of education and UK experts, can build locally appropriate assessment systems that are sustainable both in that context and within the broader relationships developed with UK institutions.

If we think about assessment as a process or system, of which an individual test forms one part, and if we conceptualise cultural relations as being ‘a concept embedded in practices’ (British Council and Goethe Institute, 2018, p. 8) then a cumulative positive impact can result from a series of interactions and activities. Seen in this way, we can begin to understand how assessment can be a very strong and powerful contributor to the cultural relations agenda.

4. The British Council and Language Assessment

The British Council formally entered the world of English language testing in 1941 when an agreement was signed between the organisation and the University of Cambridge Local Examination Syndicate (UCLES, now Cambridge Assessment English). As part of that agreement, the British Council brought its already established expertise in English Language teaching to advise UCLES on the development of its English language examinations. The importance of this partnership at that time was that the Cambridge Proficiency in English (launched in 1913) was based on contemporary theories and practices in language education; the existing expertise within the British Council in the areas of language teaching and learning was felt to offer an ideal partner to the Cambridge group. Together, the two organisations formalised what was to become an internationally recognised and distinctly UK approach to language testing – an approach that continues to influence thinking across the world.

From the outset, the contribution of the British Council was significant. In addition to providing guidance on test content, the British Council team brought a measure of innovation to the testing process not seen in UK tests before this time. From the 1970s on, the team led on the development of communicative language tests, on tests of English for specific purposes, on professionalisation of English language testing practice globally, and on theoretical conceptualisations of language assessment. A historical overview of this work is provided by Weir and O'Sullivan (2017).

The professionalisation of language assessment began with the organisation's work with UCLES in the 1940s and continued over the decades, highlighted by the formation of the English Language Consultancies Department (ELCD) in 1976; renamed the English Language Services Department (ELSD) in 1978. Ideas within this Department led to the in-house development and launch of the original English Language Testing Service (ELTS) test in 1980. Following an extensive revision project, this went on to become the highly influential IELTS test in 1990.

The Assessment Research Group (ARG) was formed by the British Council in 2012 and consists of assessment professionals with expertise in language testing theory and practice. The main objective of the ARG is to strengthen the position of the British Council and the UK as global leaders and product innovators in the field of English language assessment. The group has done this by:

- helping to diversify the British Council's portfolio of activities from test delivery to test development – the British Council's Aptis English language test was developed by the ARG
- providing academic expertise to existing and new interventions, for example, in the localisation agenda that supports the use of the British Council's Aptis test of English proficiency in a range of countries and for a range of purposes
- increasing the presence of the British Council and the UK in general on the global assessment stage
- building relationships with the UK and international assessment sectors through conference presentations, publications, contributions to the international academic community and advisory work with local ministries and institutions.

As we can see from above, the British Council's relationship with testing and assessment is long, and its expertise broad. As a result, the organisation's activities in this area are diverse and widespread. These activities can be summarised in the following ways.

Test owner and developer – British Council owns a third share of the IELTS (International English Language Testing System) test with Cambridge Assessment English and IDP IELTS Australia, as well as its own English language proficiency test, Aptis.

Test distribution and delivery – As well as delivering UK and international tests and exams for both the British Council and UK Awarding Bodies in over 100 countries across the world, the British Council has a global team to provide consultancy services to institutions and examining boards who wish to innovate their exam delivery in markets all over the world.

English Language Teaching – British Council has Teaching Centres in 50 countries with English being taught to young learners, adults and business professionals. There are also classes for exam preparation and courses for teacher development, such as CELTA and DELTA. As part of their teaching remit, teachers are involved in placement testing and diagnostic assessment, as well as formative assessment.

Supporting education reform – Through internal expertise and collaborative partnerships, the British Council supports national and institutional systemic reform to strengthen the teaching and learning of English globally.

Assessment research – Through the Assessment Research Group (ARG), the British Council is committed to research into testing and assessment to improve, innovate and push the boundaries of English language testing and assessment. This work is carried out both in-house and in collaboration with UK and international institutions

Creating opportunities – The British Council curates and participates in national, regional and international assessment conferences, webinars and workshops. This contributes significantly to the professionalisation of the field while offering UK and international researchers an opportunity to disseminate their research findings.

This indicates that the activities range from involvement with individual learners who are interested in a test to institutions who want to engage in language test development, and collaboration to governments and ministries who want systemic change and reform. The opportunities for engagement and cultural exchange are abundant and varied, and exist at all levels from individual to national.

In the British Council context, testing is sometimes perceived as an unlikely contributor to the cultural relations agenda. This may be due to the complex nature of cultural relations and the lack of a shared understanding of what testing and assessment in its entirety entails for the British Council as an organisation. It may also be due to the fact that, until very recently, the activities of the Exams division have been largely measured through reach and income rather than impact. However, the UK Government's British Council Tailored Review (Foreign & Commonwealth Office, 2019) has been a welcome catalyst for change in this respect. An important focus of the Review was the impact and, more specifically, the cultural relations impact of the work of the British Council. This meant changing the way we look at our work in Examinations; adding to the focus of reach and income to seeing our work through the eyes of our stakeholders; so, rather than looking through a single lens, seeing our work through a prism offering different perspectives.

This becomes particularly relevant when the diversity of products and services offered is considered.

In the following section, we set out how the work in testing and assessment in the British Council has operationalised the cultural relations construct. We support this with evidence from the organisation's activities in the area in recent decades.

5. Linking the British Council's Cultural Relations Mission with its Assessment Activities

The British Council has been involved with many projects over decades which have had a lasting impression on countless individuals, while influencing national and regional assessment practice (see O'Sullivan & Weir, 2017; O'Sullivan & Patel, 2019). We will re-visit a small number of these in this section, focusing on their cultural relations impact. However, before doing so, it is first necessary to reflect on how the British Council's Exams Division defines its cultural relations mission and to consider how the outcomes of these activities might be evaluated. The Exams Division defines its cultural mission in terms of the following three elements.

BUILD TRUST: The nature of UK qualifications and assessments and the way that we administer them enhances trust between the UK and other countries.

CREATE OPPORTUNITIES: The UK qualifications and assessment we provide, because of their integrity and wide recognition, have the power to change people's lives, enabling them to access life, study or work opportunities overseas or in their own countries.

DRIVE CHANGE: Qualifications and assessments are one of the most powerful drivers of improvements in teaching, learning and professional practice. They are essential to the learning process. As well as providing the exams, we often provide support for schools, organisations and individuals.

As we saw in the definition presented in the Goethe Institute and British Council (2018, p. 7) publication quoted above, the outcomes of these three goals are building:

1. greater connectivity
2. better mutual understanding
3. more and deeper relationships
4. mutually beneficial transactions
5. enhanced sustainable dialogue.

In the case studies that follow, we look to these definitions to form the basis of a systematic review in order to highlight how the British Council's assessment activities are linked to its cultural relations mission and the outcomes listed above.

5.1 The China Case Study

For many years, the British Council has delivered IELTS in China with its governmental partner, the National Education and Examinations Authority (NEEA). Since 2013, the engagement between the British Council China, the ARG and NEEA has built a strong level of trust, which, in turn, has resulted in a strong and influential partnership. In 2016, a government-to-government Memorandum of Understanding for the education sector was signed between Britain and China.

As a result of this, NEEA invited the British Council to work with them to develop the theoretical and practical procedures required to demonstrate a meaningful link between national and international examinations and the newly created China's Standards of English (CSE). As part of its mission to foster relationships between the UK and international institutions, the British Council invited Cambridge Assessment English to participate in the research and development work that formed part of the project. The entire CSE/linking initiative was a collaborative effort between the British Council and NEEA, bringing UK research and methodology expertise to the CSE project. The final report outlines in detail the process and findings of the project (Dunlea et al., 2019).

In 2019–20, the first jointly-funded research grants scheme was launched by the British Council and NEEA. This scheme is designed to offer a platform for emerging and experienced researchers in the UK and China to undertake, either collaboratively or independently, research into China's Standards of English. It is expected that fresh research partnerships will emerge, strengthening the growing trust between the two nations. In addition, Professor Barry O'Sullivan of the British Council was awarded an Advisory Professorship by Shanghai Jiao Tong University, the first person from the Arts and Humanities area to be honoured.

Table 1: Linking the China Case Study to the British Council Cultural Relations Mission

	Building Trust	Creating Opportunities	Driving Change
Cultural Relations outcomes	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue 	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue 	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue
Rationale	<ol style="list-style-type: none"> 1. At the organisational and individual level. 2. Building trust that the UK approach to research is compatible with that in China. 3. Through collaboration on different projects over several years, NEEA recognises the British Council and the broader UK as a key research and development partner. 5. By sharing expertise between the UK and China and the national level – evidenced by later joint activities. 	<ol style="list-style-type: none"> 1. Between the UK and China, particularly for Cambridge Assessment English and potentially for other UK exam boards – pathway for future recognition of their products. 4. UK tests linked to CSE offers access into the China market while recognising CSE benefits China. 5. Between NEEA, Chinese higher education institutions and individuals and counterparts in the UK. 	<ol style="list-style-type: none"> 4. By helping NEEA to recognise the level of evidence required to establish claims of links to the CSE both sides benefit. 5. By sharing expertise between the UK and China and the national level – evidenced by later joint activities.

5.2 The New Directions Case Study

New Directions is a series of high-quality conferences bringing together key officials in Ministries of Education in East Asia, representatives of local and regional testing associations and leading academics. The ARG, working with local and regional teams, shape the conferences to meet relevant needs, while at the same time showcasing UK expertise in the field. The first conference in Beijing in 2013 was instrumental in enhancing the relationship between the British Council and NEEA, while subsequent conferences have been in Tokyo, Seoul, Hanoi, Shanghai (organised in partnership with NEEA), Malaysia and Yokohama (British Council, 2020a). Due to the COVID-19 pandemic, the 2020 conference, planned for Singapore was postponed to 2021 and replaced by a full-scale online conference in October 2020.

From 2019, an additional series of New Directions was launched, this time for Latin America. The first of these conferences was held in Mexico City (2019), while the second, planned for Bogota, Colombia (2020) was postponed and replaced by a series of online events (British Council, 2020b). In all of these cases, a very positive impact has been reported as a result of the conference (Saunders and McGovern, 2019).

The engagement during these conferences and the continuity of the series has:

- strengthened relationships between the British Council and local respective Education Ministries
- provided a forum for stakeholders in East Asia to interact and engage with language assessment experts from the UK
- created an opportunity for stakeholders in East Asia to share best practice and lessons learned from within the region and elsewhere
- provided opportunities to showcase UK universities, exam boards and consultants throughout the East Asia region
- created opportunities for new collaboration between East Asia and the UK.

The British Council is now seen by Ministries of Education in East Asia as the convening authority for language assessment in the region.

As a result of the success of New Directions East Asia, New Directions Latin America started in Mexico in 2019. Saunders and McGovern's (2019) evaluation of the impact of the New Directions conference series included interviews with a range of stakeholders in the region and in the UK who reported the benefits of the British Council's convening power for relationship-building and collaboration. An example of how the conference can influence policy emerged recently when it was acknowledged that two leading academics (and policy influencers) advised their ministry on a change in direction for a major reform program based on what they had learnt at the conference (personal communication with local British Council Country Director).

Table 2: Linking the New Directions Case Study to the British Council Cultural Relations Mission

	Building Trust	Creating Opportunities	Driving Change
Cultural Relations outcomes	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue 	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue 	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue
Rationale	<ol style="list-style-type: none"> 1. Building connections between the UK and regional policy-makers and influencers. 2. Provision of opportunity to interact at all these levels. 5. Over the period, building long-term connections between people. 	<ol style="list-style-type: none"> 1. For UK-based test developers by giving them an opportunity to build connections into the East Asia market. 2 & 3. Participation in New Directions helps legitimise UK test developers and academics. 	<ol style="list-style-type: none"> 2. By exposing policy-makers and influencers to the most up-to-date thinking – e.g. it was recently acknowledged that two leading academics (and policy influencers) advised their ministry on a change in direction for a major reform programme based on what they had learnt at the conference (personal communication with local British Council Country Director). 4. Through connections brokered by the British Council at the New Directions event.

'Over time, data collection at frequent intervals will allow us to establish trends and help us to understand the impact of English language capability on national development.'

5.3 The English Impact Case Study

English Impact is a British Council global research initiative that aims to provide credible evidence on national capability in English language proficiency to inform education system development and to diagnose the impact policy decisions can have on the development of national learning systems. Over time, data collection at frequent intervals will allow us to establish trends and help us to understand the impact of English language capability on national development.

The project cuts across many areas including English language learning and assessment, as well as language education policy, evaluation and planning. Thus, far, the project has been implemented in Spain, Colombia, Bangladesh and Sri Lanka (see Shepherd and Ainsworth 2017, 2018a, 2018b).

In a number of cases, the project has led to significant changes to, or support for, language policy practices in the target country. For example, the local government in Bogota, Colombia reinforced its English as a medium of instruction agenda, based on the findings of the study.

The government of Madrid found concrete evidence of the success of their bilingual program, which in turn enabled them to rebuff criticism (often based on hearsay or on anecdotal evidence) and strengthen their policy. The fact that the British Council was chosen by the governments of these countries and regions to undertake extremely sensitive research in which their education systems risked significant criticism is clear evidence of the trust placed in the British Council. Further evidence of the impact of this project is the fact that it is, at the time of writing, being rolled out across a number of countries in Europe and East Asia.

In Table 3, we focus on the Madrid English Impact project in order to highlight its cultural relations value. The project in Madrid was undertaken in a situation where the local Ministry of Education (MoE) had received widespread criticism of its bilingual policy, with social and commercial media criticising its reach (claiming it was elitist) and ultimate value. The MoE agreed to undertake the project and publish the outcomes regardless of its findings. As it happened, these findings indicated that the criticisms were unfounded, and the policy was roundly supported by the empirical evidence provided.

Table 3: Linking the English Impact Case Study (Madrid) to the British Council Cultural Relations Mission

	Building Trust	Creating Opportunities	Driving Change
Cultural Relations outcomes	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue 	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue 	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue
Rationale	<ol style="list-style-type: none"> 2. Working together on the delivery and dissemination of the project results brought about increased understanding in both directions. On the MoE's side, this was of the expertise of the British Council in delivering such a complex project; on the BC's part, of the considerable knowledge and expertise within the MoE that strengthened the delivery of the project (and the transparency of reporting accepted). 3. The project has strengthened the relationship between the MoE and the BC due to the trust placed on each partner by each partner. This was manifest in the transparency and cooperation across the BC and MoE teams. 5. The success of the project (in part due to the positive experience in delivering the tests and partially due to the fact that the outcomes indicated that the MoE's bilingual policy was having a very positive impact on English language levels) has led to continued dialogue between the teams in the BC and the MoE in support of the way in which the bilingual policy is being implemented. 	<ol style="list-style-type: none"> 1. The project has led to improved connections to Regional MoEs. It has also led to solid research connections between local academics and academics in the UK (Universities of Bath, UK and Alcala, Spain). 3. In particular, the project has led to ongoing research within these institutions. The main connections formed are between Dr Janina Iwaniec at the University of Bath and Dr Ana Halbach of the University of Alcala. 4. The above relationship has attracted funding for two projects directly connected to the English Impact Project. 	<p>The project has done more to support change than to actually drive change – though the continued expansion of the bilingual policy has been supported to a significant degree by the project's findings. It can be argued that the combination of all five factors have led to this situation.</p>

'The Exams division has contributed significantly to the ongoing cultural relations work of the British Council.'

5.4 The Singapore Case Study

SkillsFuture Singapore is a government agency tasked with providing Singaporeans with the opportunities to develop their fullest potential through lifelong learning. In 2019 the British Council teamed up with SkillsFuture Singapore to develop the new Workplace Literacy and Numeracy (WPLN) Assessments (British Council, 2020c). Through this project, the British Council will contribute to the development of English language and workplace numeracy skills in Singapore at scale, while at the same time contributing to inclusive economic growth, by providing the assessment that will facilitate wider access to training and employment opportunities for Singaporeans. A recent development in this area is the new Singapore Government 'SG United Jobs and Skills' training package which was launched in the summer of 2020. This package includes increased funding (including living allowances while undertaking skills training) for low-skill workers to access basic literacy, numeracy and digital literacy training. WPLN scores will be used as entry and exit trackers in this project.

For the WPLN development project, the British Council Global Assessments team members (with support from the ARG in London and the Singapore Exams Operations team) created a new suite of computer-delivered workplace literacy and numeracy assessments. There are six tests in the suite: Reading, Listening, Speaking, Writing, Grammar and Vocabulary, and Numeracy. To ensure that the test was appropriate for the particular context in Singapore, the team localised the content, for example, by recording the entire listening and speaking item bank with Singaporean voice talent, in addition to the piloting, pretesting and delivery of a new Singapore-focused numeracy test. The team also benchmarked the Literacy and Numeracy tests against the existing WPLN framework to ensure continuity of acceptance of WPLN scores. In addition to the assessment development work, British Council Exams operations experts have overseen the development of a bespoke online registration, payment and reporting portal for the WPLN assessment system, and the British Council Singapore's Exams operations and customer service teams have taken responsibility for the day-to-day delivery of the assessments.

Scores from WPLN assessments are used in Singapore by people without formal educational qualifications to certify their English language and Numeracy ability. Scores are accepted for access to workplace skills qualifications courses, for entry to the Singapore Institute of Technical Education (ITE) and for qualification for a variety of jobs, including taxi drivers, security guards, and real estate agents. Test registration began in December 2019 and the test went live on 2 January 2020.

Table 4: Linking the Singapore Case Study to the British Council Cultural Relations Mission

	Building Trust	Creating Opportunities	Driving Change
Cultural Relations outcomes	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue 	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue 	<ol style="list-style-type: none"> 1. greater connectivity 2. better mutual understanding 3. more and deeper relationships 4. mutually beneficial transactions 5. enhanced sustainable dialogue
Rationale	<ol style="list-style-type: none"> 1. In this project, the British Council has built a strong partnership with the SkillsFuture (SSG) group to jointly develop the WPLN tests. 2. In working closely with SSG, both organisations have developed an understanding of each other's professional capabilities and of the contexts in which both organisations work. 3. This work has resulted in the development of a broad range of relationships across the organisations – from the technical (testing and IT) level to the delivery and policy levels. 4. The project has been mutually beneficial in that the Ministry in Singapore has gained a modern and valid set of literacy and numeracy tests while the British Council has gained from the experience of localising the literacy test and of developing, for the first time, a numeracy test. 	<ol style="list-style-type: none"> 4. The WPLN tests offer to Singaporean citizens and migrants in Singapore an opportunity to demonstrate their literacy and numeracy skills in a meaningful way; thus, opening up the possibility of meaningful and rewarding employment. They broaden the BC's influence beyond its traditional 'leaders of tomorrow' focus. 5. The British Council's Global Assessments Team and Assessment Research Group (ARG) have maintained an open dialogue with SSG to monitor both the WPLN tests (ensuring that they are functioning as planned) and the impact the tests are having on stakeholders. Both of these are seen by the British Council and SSG as being critical to the long-term viability of the tests. 	<ol style="list-style-type: none"> 2. The discussions relating to monitoring and evaluation of the WPLN tests, together with the transparency of the development process, have prompted the partners (British Council and SSG) to look critically at how the tests are functioning now that they are operational. This approach is expected to contribute to the future direction taken by SSG with regards to the WPLN and other tests. 5. This work has paved the way for a sustainable dialogue between the partners on how testing and assessment within this Singaporean context might look in the medium and longer terms.

As well as the SSG relationships, the British Council has built strong relationships with the 'Special Schools' in Singapore as part of its long-term commitment to equality, diversity and inclusion (EDI) as laid out in its EDI policy and strategy (2017a, 2017b). Students in those schools transition to technical education and WPLN scores are used in place of O or N Levels. The key relationships here are with Heads of Curriculum, with the local British Council team investing considerable time in working with them on the test format and construct. The BC operations team also works closely with school administrators to deliver a range of accommodations for their candidates (these include special measures, such as additional time or Braille test papers, to ensure that candidates are able to demonstrate their best level of ability despite any disability). These initiatives have resulted in very positive feedback on the processes around delivery and accommodations for the special school candidates.

The project is strong evidence of the growing reputation in the region of the British Council as a leader in the area of educational assessment.

5.5 Commentary

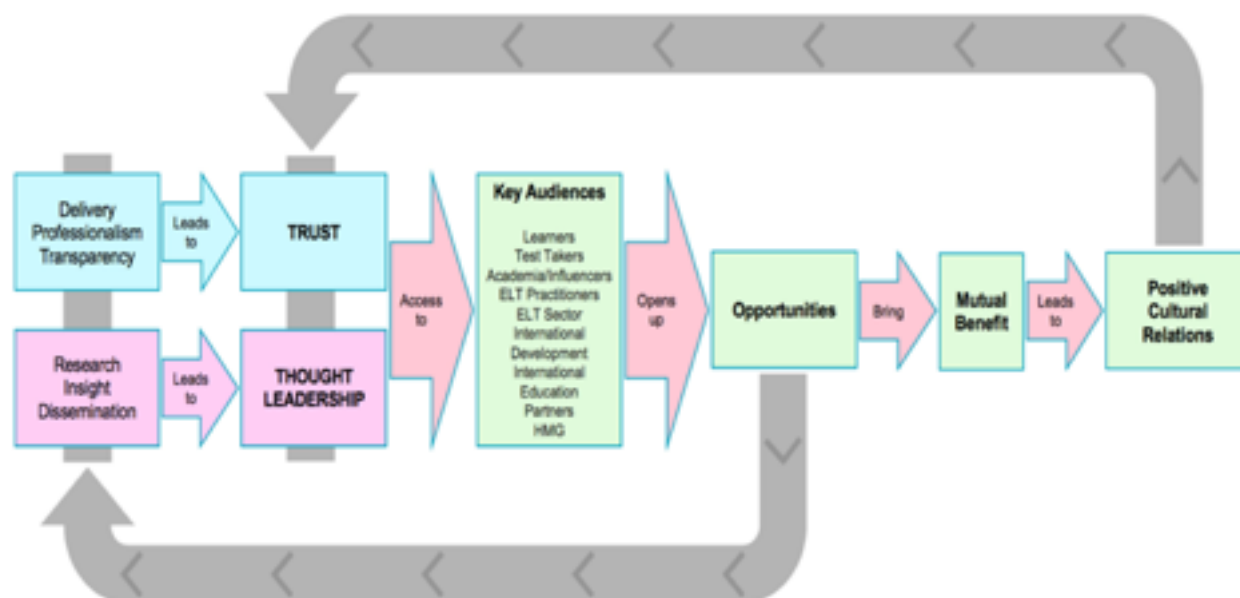
The evidence presented in these case studies suggests that the Exams Division has contributed significantly to the ongoing cultural relations work of the British Council. These four recent (or ongoing) projects should be seen in the context of the many projects undertaken or supported by the British Council reported in the Weir and O'Sullivan (2017) overview of the organisation's activities over many decades in the area of language test research and development (see Chapter 2 of that book). The evidence presented, however, is observational in nature, and while the above summaries have been verified by teams on the ground in each case, there remain some question over the strength of the claims we make. While we stand over these claims, they would clearly be strengthened significantly through the addition of systematic, empirically derived data. In the following section, we suggest a practically operationalisable model for generating such data in a systematic and sustainable way.

6. Building Our Evidence: An Integrated Approach

In this section, we build on the arguments put forward above about how our work in assessment is central to the British Council's cultural relations remit. We make the case for an integrated evidence strategy to demonstrate how, and with whom, our work in assessment builds connections, understanding and trust, and the consequences of this. Our central argument is that to be successful in this endeavour, we will need to draw upon high quality monitoring of outcomes, well-designed program evaluations, and purposefully designed research and analysis. These need to be established and implemented as a single evidence strategy.

Figure 1 indicates graphically how the British Council sees the process of trust building and thought leadership as forming a key part of a longitudinal and iterative process. Projects such as those described here build the trust of prospective partner organisations and institutions through timely and efficient project delivery, as well as professional expertise and the transparency of the process, often engaging partners in the development, delivery and reporting work. This trust allows the organisation access to key audiences which, in turn, leads to the sort of mutual benefit which characterises positive cultural relations, which in turn leads to increased trust and so on. In the same way, our research, insight and dissemination activities lead to the organisation being recognised as a thought leader in the field of language testing. This also leads to access to key stakeholders (though usually in a different way and to different groups), which leads to opportunities. These opportunities include the kind of activities that drive the research and insight strategies while at the same time offering the same mutual benefits that lead to positive cultural relations.

Figure 1: *The iterative nature of trust and thought leadership building*



6.1 Our Cultural Relations Evidence Needs

Three of the recommendations from the 2019 Tailored Review (Foreign & Commonwealth Office, 2019) emphasised the need for the British Council to invest in stronger, systematic evidence of the impact of its work, and the better use of evidence to guide resource investments. Doing this requires clear outcome frameworks, rigour in method design, analysis, and use, and an ‘evidence literate’ organisational culture, that is able to interpret and apply both positive and negative results. The case studies outlined in the previous section present a strong rationale as to how British Council’s work in English assessment can have cultural relations impact; our need now is to develop greater rigour in evidencing this relationship, showing the changes that have been brought about, the mechanisms through which change occurs, and learning as to how to maximise impact.

6.2 The Types of Evidence We Need for Cultural Relations

There are three related but distinct approaches to building our evidence:

1. **Monitoring** of outcomes against a pre-designed framework that describes the outcomes being sought, their definitions and a process for tracking them prospectively that adheres to international standards of quality (by way of example, see <https://www.measureevaluation.org/our-work/data-quality>). High quality outcome monitoring can be in relation to a single program or service, or range of programs and services that operate towards a similar end. It is not a substitute for comprehensive program evaluation, but is a source of critical, evaluative evidence.
2. Well designed, **independent evaluations** that draw upon multiple data sources to make evaluative judgments on the extent to which a program or area of work has contributed to a set of outcomes, ideally with some assessment of a counter-factual (Quinn Patton, 2018).

Evaluations can be conducted at the program, portfolio or country level. The quality of evaluations is, in part, dependent upon the quality of outcome monitoring. Evaluations should focus on both program implementation and program outcomes.

3. Purposefully designed and delivered **research and analysis** that contributes new generalisable knowledge and understanding, but which does not necessarily have a single program focus and which will often be to test or generate hypotheses, rather than seeking to comprehensively evaluate the implementation of a program. Such research and analysis will often have its origins in high quality outcome monitoring which can be ‘hypothesis generating’, but which require well designed and implemented research to test.

These three approaches to evidence building are conceptually distinct but overlap and are inter-related. Program evaluation or operations-focused research will, in most cases, require high quality outcome monitoring in order to reach well-founded conclusions about program, portfolio or country level impact, or about a specific hypothesis. Research to test hypotheses can provide the theoretical basis for a program, area of work, or method of delivery.

The approach can be summarised as shown in Figure 2. The whole evaluation starts by defining its objectives and then clearly outlining these in the Outcomes Framework. Based on this framework, the monitoring phase begins at the outset of program or service delivery and continues throughout. If this fails to provide evidence of success (as outlined in the Outcomes Framework) then program evaluability is likely low. An independent impact evaluation is unlikely to be of value in this scenario.

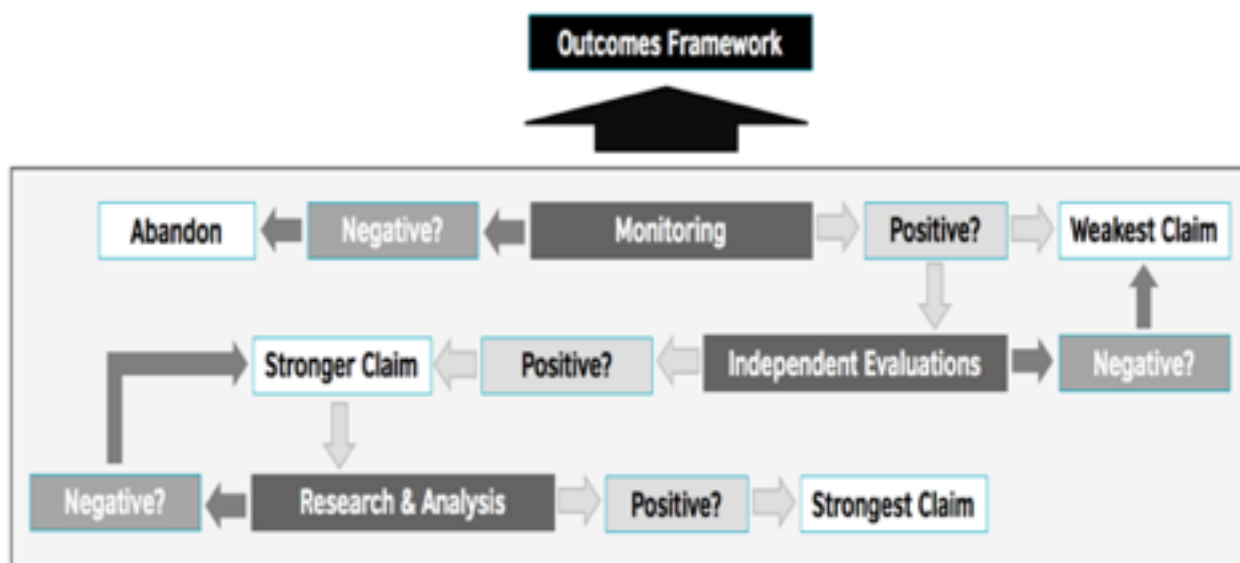
If monitoring data does produce evidence of success, we progress to the Independent Evaluations phase. If this fails to provide any additional meaningful evidence, then at best we can only claim the weakest level of claim – we must accept that evidence from this phase may not support the claim made following the monitoring phase and may result in a full or partial rejection of that claim or set of claims. If appropriate evidence is forthcoming, we can argue for a stronger claim and proceed to the Research and Analysis phase.

On many occasions, it may be deemed enough for a project to make these stronger claims – e.g., as neither time nor resources will have been allocated to allow for a longitudinal study or series of studies.

In the final phase, a series of research projects are initiated in order to further explore specific questions related to the cultural relations impact of the initiative, project or event. Typically, these projects are undertaken by external research teams and provide additional evidence in support (or not) of the claims made. Where the evidence proves positive, the project team can justify the strongest level of support for their claims. As with the previous stage, if this final phase either fails to materialise or fails to provide additional appropriate evidence, we are left with the stronger claim – though again, if a research agenda fails to support (or contradicts) the claims made earlier, we must be prepared to either revisit these claims or abandon them entirely.

The main issue with this approach is of knowing when enough is enough. This same issue has plagued educational assessment theorists for many years – specifically in relation to the argument put forward by Messick (1989 and elsewhere) that validation was an ongoing process with no clear end point. On the other hand, Kane (2006 and elsewhere) argues that his Toulmin-argument based approach to validation provides a way in which the developer can draw a line under the validation research undertaken, essentially closing off the argument. This approach provides an interesting avenue to pursue with regard to the required breadth of impact arguments.

Figure 2: *The types of evidence pathway*



6.3 The Need for an Integrated Approach to Cultural Relations Evidence

Delivering evidence of the cultural relations impact in English language assessment requires an integrated strategy involving all three of these approaches. We need outcome frameworks that establish standardised, prospective monitoring protocols across different areas of work. This will be of value in itself and will also provide the platform to guide evaluation, research and analysis.

We need a well-governed and appropriately-funded evaluation program that identifies the key programs and portfolios where investment in evaluation would add to our evidence base and learning. One example is the WPLN program in Singapore where we are taking a comprehensive approach to evaluating the impact of the program on candidates, test preparation procedures, and policy debates.

We need research and analysis to answer both wider hypotheses, such as the meta review and analysis of the relationship between English language proficiency and engagement with the cultural, educational, and business assets that the UK offers, conducted in 2019 (Culligan & Pelletier, 2019).

To do this will require a clearly defined evidence agenda and strategy for cultural relations impact for our Exams work, drawing upon the different forms of evidence outlined above.

It will also require delivery teams in our country offices who understand cultural relations outcomes, the relevance to their work and how they contribute to the organisational purpose. It will also require projects to have high quality monitoring, evaluation and research plans developed as part of wider implementation plans.

As we set out in this paper, cultural relations impact in the area of English language assessment has multiple impact pathways. These can be described as:

- Wider knowledge of the English language that contributes to greater engagement with UK assets. Language assessment plays a critical role in language proficiency.
- Our expertise in ELA acting as a high-value soft power asset, making us attractive to overseas governments and institutions as strategic partners in their ambitions around English teaching, learning and assessment.
- Leveraging our soft power assets to create opportunities for the UK ELA sector. While British Council has expertise, we are also key to opening up opportunities for wider UK ELA sector engagement overseas.

Building our evidence around these pathways requires an integrated approach that develops our learning and allows us to confidently articulate the ways in which our work in Exams delivers impact that is central to our remit as a cultural relations organisation.

The other dimension that we need to be mindful of is time. Connections, understanding and trust are iterative and have no defined end point. They are properties that develop, and are maintained, over long time periods. They are the product of multiple forms of engagement, and they can both increase and decrease. Our evidence strategy for Exams needs to take both a short and long-term perspective, identifying both what is going to be useful now, but also how it will contribute to a longer-term understanding of how our work in this area delivers cultural relations impact.

'Connections, understanding and trust are iterative and have no defined end point. They are properties that develop, and are maintained, over long time periods. They are the product of multiple forms of engagement, and they can both increase and decrease.'



7. Conclusion

In this paper, we have presented a view shared across the British Council on the nature of Cultural Relations. We have shown how the work in language testing and assessment carried out within the organisation can be seen to contribute to its cultural relations mission. We presented this in terms of the interaction between the Exams Division's vision of its cultural mission (building trust, creating opportunities and driving change) and the broader organisational goals as outlined in the joint Goethe Institute and British Council report (2018).

In this regard, we presented evidence which was supported by local British Council staff who had participated in a series of very different projects across the world. It was clear from this section of the paper that there is certainly the basis for a claim that the projects referred to meet the criteria for positive cultural relations impact as defined within the organisation. However, it is clear that the strength of this claim is limited by the kind of evidence presented. In order to strengthen any such claim, we then propose a more robust, integrated approach.

In this approach, we outline three distinct phases of evidence gathering: monitoring, independent evaluations, and research and analysis. While the evidence presented here in the four case studies can be seen as being appropriate to the first of these, the fact that the relevant areas were not identified *a priori* weakens the argument that they represent a strong claim for cultural relations impact. We would therefore argue that while the framework identified in the case study tables may offer an appropriate approach, unless the evaluator sets out to complete them before the project begins in earnest, they cannot serve as the basis for a convincing argument.

In order to strengthen the claim of appropriate positive cultural relations, it is important to carry out formal, structured evaluations. This requires all program proposers to build a budget into their proposals to allow for such an evaluation. They should also ensure that the work is undertaken by independent experts with knowledge of the program and of the context. This latter expertise is often missing where program evaluations are typically undertaken without sufficient recognition of the vagaries of specific contexts where global frameworks may (or may not) be appropriate.

Finally, we argue that to attain the highest level of claim neither the program nor the evaluation should not be seen as a singular activity. Instead, the organisation should set about developing a research agenda built around developing new evidence-based evaluation approaches that can be systematically localised for use in specific contexts. Of course, this has significant resource implications in terms of cost and time and may not always be feasible. However, where possible, this level of evidence should be the ultimate goal of all impact studies.

It is also clear that the three areas identified by the British Council Exams Division as defining its cultural mission may not be fully addressing the issues. From the evidence and arguments presented in this paper, it is clear that an additional area should be added – this relates to the way in which we support English language learning and assessment. This reference to the everyday activities of the organisation (in fact any organisation) is critical in locating the cultural relations impact of its work on specific initiatives, projects or events within the reality of its core activities.

We readily acknowledge that the integrated approach we propose here is unlikely to offer fast and easily resourced evidence of cultural relations impact. We recognise also that the strength of the claims we make about different projects or programs will not always need to span all three integrated levels. However, it appears clear to us that the challenge to the British Council, and to other organisations engaged in cultural relations activities, is to identify the extent and range of evidence required and the impact of this on the strength of the associated claim of a positive impact.

Cultural relations should not be seen as a short-term outcome of a specific project or program. As can be seen in Figure 1, the path to positive cultural relations outcomes is typically long and complex. Viewing an evaluation approach as appropriate to the short or even medium term is likely to result in disappointment.

'In this paper, we have shown how our work in language testing and assessment carried out within the organisation can be seen to contribute to its cultural relations mission.'

References

- British Council (2017a). Equality Policy. Accessed from: <https://www.britishcouncil.org/sites/default/files/equality-policy-2017.pdf>
- British Council (2017b). Equality, Diversity and Inclusion Strategy. Accessed from: <https://www.britishcouncil.org/sites/default/files/edi-strategy-2017.pdf>
- British Council. (2020a). New Directions in English Language Assessment: East Asia. Website accessed from: <https://americas.britishcouncil.org/new-directions>
- British Council. (2020b). New Directions in English Language Assessment: Latin America. Website accessed from: <https://www.britishcouncil.sg/new-directions/about>
- British Council. (2020c). Workplace Literacy and Numeracy (WPLN) Assessments. Website, accessed from: <https://www.britishcouncil.sg/exam/wpln>
- Culligan K, Pelletier, A. (2019). Cultural Relations, Evidence of Impact. British Council internal document.
- Dunlea, J., Spiby, R., Wu, S., Zhang, J., & Cheng, M. (2019). China's Standards of English Language Ability (CSE): Linking UK Exams to the CSE. Technical Report on Linking UK Exams to the CSE. British Council Validation Series, VS/2019/003. Accessed from: https://www.britishcouncil.org/sites/default/files/linking_cse_to_uk_exams_5_0.pdf
- Fairbank, J. K. (1968). Introduction. In J. K. Fairbank (ed.), *The Chinese world order: Traditional China's foreign relations*. Cambridge MA: Harvard University Press.
- Foreign and Commonwealth Office. (2019). British Council Tailored Review. Accessed from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/785297/British_Council_Tailored_Review.pdf
- Goethe Institute and British Council (2018). Cultural Value. Cultural Relations in Societies in Transition: A Literature Review. Accessed from: https://www.britishcouncil.org/sites/default/files/lit_review_short_working_paper_final_final.pdf
- Kane, M. T. (2006). Validation. In R. Brennan (Ed.), *Educational measurement*, 4th ed. (pp. 17–64), Westport, CT: American Council on Education and Praeger.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 13–103). New York: American Council on Education & Macmillan.
- O'Sullivan, B., & Patel, M. (2020). English Language Assessment as Cultural Relations. In J. P. Singh (ed.), *British Council Cultural Relations Collection*. Accessed from: <https://www.britishcouncil.org/sites/default/files/english-language-assessment-cultural-relations.pdf>
- Quinn Patton, M. (2018). Evaluation Science, *American Journal of Evaluation*, Vol. 39, Issue 2: 183–200.
- Shepherd, E., & Ainsworth, V. (2017). *English Impact: An Evaluation of English Language Capability*. Madrid, Spain. London: British Council. Accessed from: https://www.britishcouncil.org/sites/default/files/g303_03_english_impact_report_madrid_web1.pdf
- Shepherd, E., & Ainsworth, V. (2018a). *English Impact: An Evaluation of English Language Capability*. Bogota, Colombia. London: British Council. Accessed from: https://www.britishcouncil.org/sites/default/files/h203_01_english_impact_report_colombia_aw4.pdf
- Shepherd, E., & Ainsworth, V. (2018b). *English Impact: An Evaluation of English Language Capability*. Sri Lanka. London: British Council. Accessed from: https://www.britishcouncil.org/sites/default/files/h203_01_english_impact_report_sri_lanka_aw5.pdf
- Singh, J. P. (2019). British Council and Cultural Relations: Betwixt Idealism and Instrumentality. In J. P. Singh (ed.), *British Council Cultural Relations Collection*, (p. 7). Accessed from: <https://www.britishcouncil.org/sites/default/files/british-council-cultural-relations-idealism-instrumentality.pdf>
- Wade, G. (translator). (2005). *Southeast Asia in the Ming Shi-lu: an open access resource*. Singapore: Asia Research Institute and the Singapore E-Press, National University of Singapore. Accessed from: <http://www.epress.nus.edu.sg/msl/reign/hong-xi/year-1-month-7-day-12>
- Weir, C. J., & O'Sullivan, B. (2017). *Assessing English on the Global Stage: The British Council and English language testing, 1941–2016*. London and Sheffield: British Council & Equinox Publishing.