

DESIGNING AN INTERNATIONAL ENGLISH LANGUAGE EVALUATION USING A STANDARDISED TEST

EXISTING LANDSCAPE

CHERRY GOUGH

**COUNTRY DIRECTOR, BRITISH COUNCIL,
VIETNAM**

THE PISA MODEL

The motivation for PISA's existence was a perceived need to fill a gap in the extensive set of indicator-based information on education systems that the OECD has been providing ...in its annual Education at a Glance reports. The information in these reports extensively covers system input and process ... Each report is considered to offer ... a rich, comparable, and up-to-date array of indicators that reflect a consensus among professionals on how to measure the current state of education internationally. (OECD, 2010a, p.17) **The gap in this information was a set of indicators with which countries might evaluate the outcomes of their educational provision in terms of their students' achievement.**

2011: Oxford University Centre for Educational Assessment, Policy Effects of Pisa

INTERNATIONAL PERSPECTIVES

More and more countries are looking beyond their own borders for evidence of the most successful and efficient policies and practices. Over the past decade, the OECD Programme for International Student Assessment (PISA) has become the world's premier yardstick for evaluating the quality, equity and efficiency of school systems.

OECD (2013), Lessons from PISA 2012 for the United States, Strong Performers and Successful Reformers in Education, OECD Publishing.
<http://dx.doi.org/10.1787/9789264207585-en>

THE PISA COMMITMENT

PISA represents a commitment by governments to monitor the outcomes of education systems through **measuring student achievement on a regular basis and within an internationally agreed common framework**. It aims to provide a new basis for policy dialogue and for collaboration in defining and implementing educational goals, in innovative ways that reflect judgments about the skills that are relevant to adult life.

OECD, (2009) PISA 2009 Assessment Framework: Key Competencies in Reading, Mathematics and Science. Paris: OECD Publishing.

FOR ENGLISH, CAN WE...?

- **evaluate the outcomes of ... educational provision in terms of ... students' achievement**
- **[find] evidence of the most successful and efficient policies and practices**
- **measure student achievement on a regular basis and within an internationally agreed common framework**

REACTIONS TO PISA RESULTS

- **‘England needs a rocket under them to improve their PISA scores’, says Minister for Education.**
- **The PISA results of 2000 and 2003 left Norway in a state of ‘PISA Shock’**
- **The response from international media [to the Shanghai results], in the USA, e.g. New York Times, The Washington Post, was swift, sharp, and even sometimes hysterical. It was considered a ‘Sputnik’ moment, a ‘Pearl Harbor Day’, and a ‘massive wake-up call’ for American education (Secretary of Education, Arne Duncan).**

RESPONSES TO PISA...

- England: [A] quandary for England is how national test results such as key stage tests and GCSEs have shown rising outcomes whilst comparisons with other countries have not changed... The coalition government has signaled ...that it values international comparisons. [However] **there are many policies contained in the [2010] White Paper that have no obvious relationship with the PISA findings.**

- Norway: a series of reforms of both curriculum and assessment. A national quality assessment system (NKVS) was introduced in 2004 which included national tests and a web-based portal (Skoleporten) for presentation of data for school evaluation.
- Switzerland: Within political circles the impact of PISA, and TIMSS ...was to accelerate a long-awaited process of harmonisation in educational structures, practices and curriculum in this federated country, with its history of cantonal autonomy in education and consequent variety of provision.
- The international surveys have been used in support of a wide-ranging review and reform of education in France. A first plan of action the Government is to 'refocus on fundamentals', by reforming the primary school curriculum, introducing a strategy to fight illiteracy, and launching an initiative in science. A second plan of action aims to introduce personalised learning assistance throughout the system, to help students in difficulty. The third plan of action is to 'personalise' educational resourcing, such as making schools autonomous managers of their own budgets.

Vietnam's performance in the last international Pisa tests was a stunning achievement.

Participating in the tests for the first time, the country's 15-year-olds scored higher in reading, maths and science than many developed countries, including the United States and the United Kingdom.

BBC

2016 RISE research: Vietnam

..it is crucial to look at the Vietnam case and understand *what worked* in their context, and *how* it worked, and what lessons we can learn that can be applied elsewhere... In short, Vietnam is an opportunity for much research to be done, both in lessons for other countries and for sustainability of the success story within Vietnam itself.

Key institutions: **University of Minnesota**, the Centre for Analysis and Forecasting, the Vietnamese Academy of Social Sciences, and the Mekong Development Research Institute.

Other affiliated institutions: the (UK) Institute for Fiscal Studies, Leiden University, University College London, the University of Oxford, and the World Bank.

CAN WE DO THE SAME FOR ENGLISH?

- EF EPI - <http://www.ef.co.uk/epi/>
- European survey of languages - <http://www.surveylang.org/>
- ETS (TOEFL and TOEIC data)
https://www.ets.org/s/toeic/pdf/ww_data_report_unlweb.pdf
- IELTS - <https://www.ielts.org/teaching-and-research/test-statistics>

THE BRITISH COUNCIL MODEL

- To evaluate English language competence of students at grade 10/age 15
- To test performance via a reliable test of proficiency in all four skills
- To produce an index where overall country scores can be used as one measure of success of the country's English language policy
- To produce a national English proficiency score which can also be correlated against other national features: e.g. economic success, innovation

And perhaps...

- To offer a means of identifying and learning from successful countries
- To create a stimulus to focused investment in EL teaching
- To increase the status of English teachers in the community.

DO WE WANT TO DO THE SAME FOR ENGLISH?

自信 zì xìn

自省 zì xǐng

METHODOLOGICAL CHALLENGES TO DELIVERY

ELIZABETH SHEPHERD

SENIOR RESEARCHER, BRITISH COUNCIL, UK

CREATING EQUIVALENCE

‘Bias refers to the presence of nuisance factors in cross-cultural research. Three types of bias are distinguished, depending on whether the nuisance factor is located at the level of the construct (construct bias), the measurement instrument as a whole (method bias) or the items (item bias).

Equivalence refers to the measurement level characteristics that apply to cross-cultural score comparisons; three types of equivalence are defined: construct (identity of constructs across cultures), measurement unit (identity of measurement unit), and scalar equivalence (identity of measurement unit and scale origin).

Bias often jeopardizes equivalence’.

(Van de Vijver 1998)

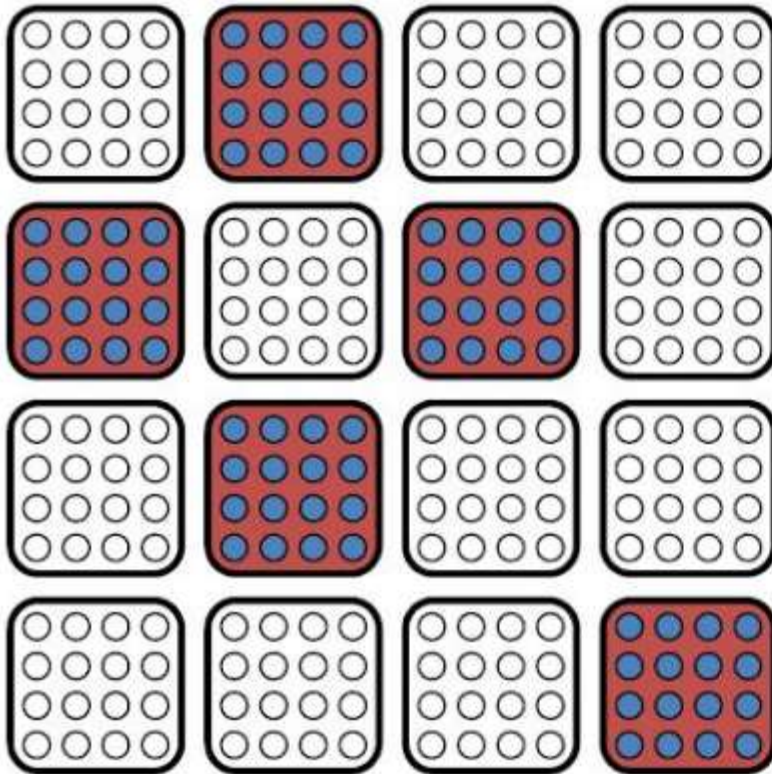
MINIMISING BIAS

- Construct bias
- Method bias
- Item bias

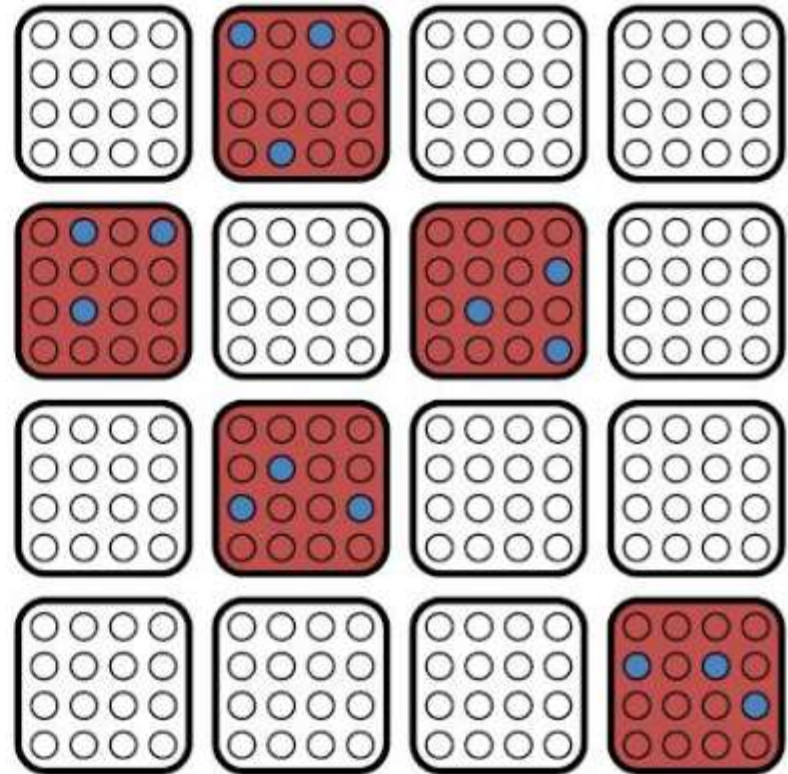
CONSTRUCT BIAS

- Students within the international comparison unit
- Enrolled in the grade that represents ten years of schooling counting from the first year of ISCED Level 1
- providing the mean age at the time of testing is at least 15.5 years
- who are currently studying A minimum of 90 minutes of formal English study per week as part of their studies at this grade level.

METHOD BIAS



1st stage cluster sampling

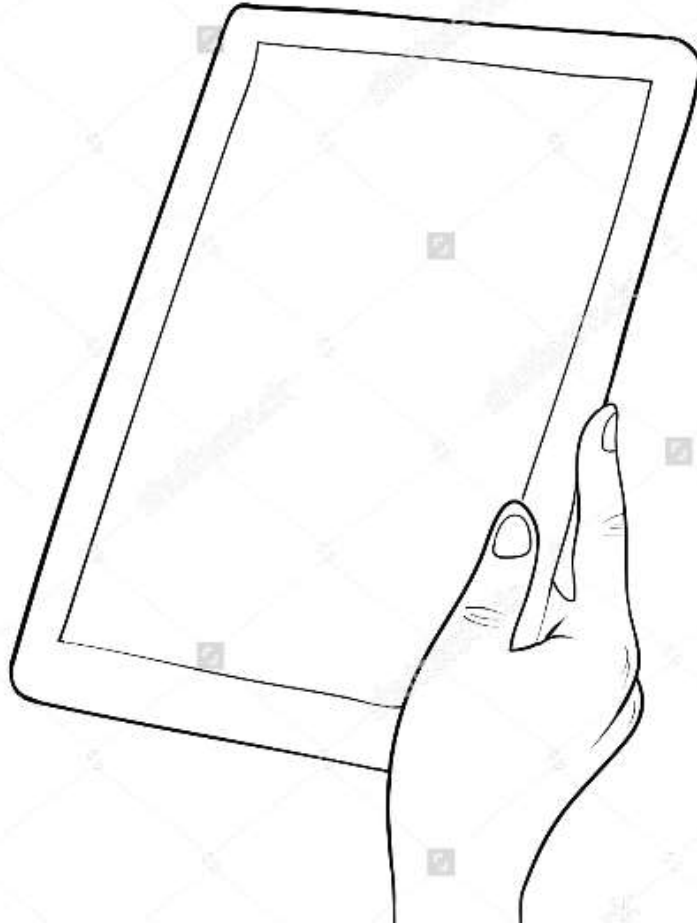


2st stage cluster sampling

METHOD BIAS

The logo for the Assessment and Quality Improvement Research (ACER) organization. It consists of the letters 'A', 'C', 'E', and 'R' in a white, serif font, each contained within a dark purple square. The squares are arranged horizontally and separated by thin white vertical lines. The 'R' square has a white corner cutout in the top right, giving it a document-like appearance.

METHOD BIAS



METHOD BIAS



eTIMSS 2019 is designed to keep pace with an increasing reliance on computer-based assessments while:

Improving measurement

- eTIMSS assesses complex areas of the TIMSS Assessment Frameworks that have been difficult to measure using paper and pencil
- eTIMSS can stimulate student motivation through its interactive and animated tasks

Increasing operational efficiency

- eTIMSS will improve item development, translation and translation verification, and data entry and scoring, while reducing printing and shipping costs

TIMSS, TIMSS Numeracy, and eTIMSS are projects of IEA (International Association for the Evaluation of Educational Achievement). Headquartered in Amsterdam, IEA has conducted international comparative studies of educational achievement since 1959.

The TIMSS projects are directed by the TIMSS & PIRLS International Study Center at Boston College. TIMSS together with PIRLS, which assesses reading, comprise IEA's core cycle of studies about achievement in three fundamental subjects—mathematics, science, and reading.

eTIMSS is jointly developed by the TIMSS & PIRLS International Study Center and the IEA Data Processing and Research Center in Hamburg.



THE ANATOMY OF eTIMSS

eTIMSS consists of a series of interconnected software modules hosted on the IEA DPC servers. With eTIMSS, National Research Coordinators can:

- Use the Item Builder to develop achievement items and Problem Solving and Inquiry tasks for the eTIMSS tablet and stylus environment
- Use the Online Translation System for translating the eTIMSS items into the language of instruction and having them verified by IEA
- Use the Tablet Player to administer the eTIMSS assessment—present the items in tablet and stylus format, record students' responses, and upload the data to the eTIMSS server
- Use the Online Data Monitor to observe the progress of the data collection
- Use the Online Scoring System to review students' written responses and score them according to the eTIMSS scoring guides

Once data collection is complete, the data are sent to the TIMSS & PIRLS International Study Center for review and analysis.

For country enrollment, contact:

Dr. Paulina Koršňáková, IEA Secretariat
p.korsnakova@iea.nl

www.iea.nl
timss.bc.edu



TIMSS & PIRLS
 International Study Center
 Lynch School of Education, Boston College

ITEM BIAS

ENGLISH

Ingles

الإنجليزية

英语

영어

WHERE WE ARE NOW

- Ministry of education collaboration
- Technical solution design and hardware testing
- Questionnaire item translation and performance trial
- Sample frame creation
- Field operations team training
- School liaison preparation
- Theory of change and evaluation planning
- Open access data sourcing

Q&A

cherry.gough@britishcouncil.org

elizabeth.shepherd@britishcouncil.org