Effectiveness:
How do learning assessments influence policy in developing countries?

A project funded by The Australian Agency for International Development (AusAID)

Presented at the Network for Education Quality Monitoring in the Asia-Pacific Region (NEQMAP) Meeting in Bangkok, 27-28 March 2013

Dr Petra Lietz

Australian Council for Educational Research
Systematic Review of the Impact of National and International Assessment Programmes on Educational Policy, Particularly Policies Regarding Resource Allocation and Teaching and Learning Practices in Developing Countries

Review team:
Maura Best, Pat Knight, Petra Lietz, Craig Lockwood (Joanna Briggs Institute), Dita Nugroho, Mollie Tobin
Overview

- Background
- Concepts
- Study questions
- Methodology
- Findings
- Final thoughts
BACKGROUND
Background

• In 1964, the International Association for the Evaluation of Educational Achievement (IEA) conducted the first internationally comparative study in mathematics; 12 countries participated.

• By 2009, about 70 per cent of the countries in the world participated in some form of regional or international assessment program.

• Since the 1980s, the number of national assessment programs has also increased, e.g. in Central and South America:
Background

International Assessment Programs:

• The International Association for the Evaluation of Educational Achievement (IEA): PIRLS (reading Grade 4), TIMSS (maths, science Grades 4&8), ICCS (Civics), ICLS (ICT literacy); since 1960; 59 systems in PIRLS/TIMSS 2011.

• Organisation for Economic Cooperation and Development (OECD): PISA (maths, reading, science, financial literacy); 15-year-olds; since 2000; 67 participating systems in 2012.
Background

Regional Assessment Programs:

- **Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ);** 15 countries; since 1995; supported by UNESCO’s IIEP

- **Latin American Laboratory for the Evaluation of Quality in Education (LLECE);** 15 countries; since 1997. LLECE sponsored a second assessment (SERCE) study in 2007. TERCE planned for 2013; supported by UNESCO’s Regional Bureau for Education in Latin America and the Caribbean.

- **The Conference of Education Ministers of Countries Using French as the Language of Communication (CONFERMEN);** conducts the Program for the Analysis of Educational Systems of States and Governments (PASEC); founded in 1960; 41 members; 24 participating.
Concept: Assessments

- National and international assessment programmes will be understood to be assessment programmes conducted in primary and secondary education, and will exclude assessment programmes that are undertaken at the sub-national level (e.g. province, district).
- References to non-standardised assessments were not considered in this review.
Concept: Policy Cycle

1. Agenda setting: Awareness of and priority given to an issue or problem

2. Policy formulation: The ways analytical and political options and strategies are constructed

3. Policy implementation: The forms and nature of policy administration and activities on the ground

4. Monitoring and policy evaluation: The nature of monitoring and evaluation of policy need, design, implementation and impact.

Simplified Model of the Policy Cycle; Source: Sutcliffe and Court (2005)
Further Concepts

• Developing countries:
  – AusAID and World Bank lists (157 combined)

• Resource allocation:
  – E.g. $ per student, instructional materials, student-teacher ratio, teacher qualifications

• Teaching and learning practices:
  – E.g. Classroom management and discipline, instructional strategies and activities, teacher collaboration
METHODOLOGY
## Framework Synthesis Table

<table>
<thead>
<tr>
<th>Policy</th>
<th>Facilitators</th>
<th>Barriers</th>
<th>Quality</th>
<th>Equity</th>
<th>Accountability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resource allocation</td>
<td>U</td>
<td>V</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching and learning</td>
<td>W</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>Y</td>
<td>Z</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Policy Stages

- Agenda setting
- Policy formulation
- Policy implementation
- Monitoring and evaluation

### Regions

- **National**
  - Primary
  - Census: A, B, C, D, E
  - Sample: F, G, H, I, J
- **Regional**
  - Secondary
  - Census: K, L, M, N, O
  - Sample: P, Q, R, S, T
- **International**
  - Both primary and secondary

### Assessment types

- National
- Regional
- International

### Levels of Education

- Primary
- Secondary
- Both primary and secondary

### Sampling approach

- Census
- Sample
Sample Keywords

• Describing assessment program:
  – Type (Sub-national, National, Regional, International)
  – Education level (Primary, Secondary, both)

• Describing policy impact:
  – Stage of the policy cycle (Agenda setting, Formulation, Implementation, Monitoring and evaluation, No impact)
  – Type of policy:
    • Resource allocation (e.g. Instructional materials, Class size, Teacher preparation)
    • Teaching and learning (e.g. Classroom management, Student-oriented pedagogy)
The Brazilian National Evaluation System of Basic Education: Context, Process, and Impact

**Abstract**
Describes the Brazilian National System of Evaluation of Basic Education and analyzes its impact on educational policies and practices, especially at the state level. Data from about 15 interviews, site visits, document review, and a survey completed by representatives of 23 state education departments show the strong points and weaknesses of the evaluation system. (SLD)

**Title**
Studies In Educational Evaluation

**Author(s)**
Crespo M; Soares Jose Francisco; de

**Publication Information**

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>2000</td>
</tr>
<tr>
<td>Pub type</td>
<td>Journal, Article</td>
</tr>
<tr>
<td>Item IDs</td>
<td>Internal: 4929417 Imported:</td>
</tr>
<tr>
<td>Journal title</td>
<td>Studies In Educational Evaluation</td>
</tr>
<tr>
<td>Short title</td>
<td>Crespo (2000)</td>
</tr>
<tr>
<td>Pages</td>
<td>105-125</td>
</tr>
<tr>
<td>Volume</td>
<td>26</td>
</tr>
<tr>
<td>Publisher</td>
<td>Institution</td>
</tr>
</tbody>
</table>
RESULTS
Included literature largely discuss countries in South America and Africa – South of Sahara, with some from Asia.
• National assessments were most frequently mentioned, largely in South America and Asia

• Regional assessment studies concentrated in Africa – South of Sahara (SAQMEC and PASEC)

• Sub-national assessment examples only found in larger countries of South America (e.g. Chile, Argentina, Brazil) and South Asia (e.g. India)
Stated Goals of Assessment

‘Other’ goals:
- Inform future assessments and build technical capacity
- Enable broad international comparisons
- Provide inputs to be used at local- or school-level
- Evaluate the effects of specific policies

% of total mentions

- Quality: 34%
- Equity: 20%
- Accountability: 19%
- Leverage: 12%
- Other: 10%
- Unknown/NA: 5%
Additionally, of the total number of studies included, 13 studies (24%) explicitly note “No Impact on Policy Process” -> 4 reasons.....
Reasons for “no impact”

• Where reports explicitly noted that a program has had *no impact* on the policy process, the key barriers were:
  – Poor quality program (not designed to facilitate meaningful comparisons and analyses, or with policy goals in mind)
  – Data from international programs cannot answer the questions/concerns of national policymakers
  – Analysis in outputs (descriptives, correlations and even IRT specifically mentioned) cannot be interpreted to policy
  – Poor dissemination to policymakers and schools
Types of policy impacted

- Resource Allocation: 91
- Teaching and Learning Practices: 35
- Curriculum Standards: 23
- OTHER: 22
- Assessment Policy: 19
- Performance Standards: 18
- Accountability Policy: 10
- Community/Parent Engagement Policy: 7
- Unknown/Not Applicable: 6

- In 65% of studies involving international assessment programs, they were found to impact curriculum standards.
- ‘Other’ policies impacted included well-being funding and interventions, performance-based pay and incentives, and expansion of early childhood education.
<table>
<thead>
<tr>
<th>Resource Allocation Decisions</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructional Materials</td>
<td>15</td>
</tr>
<tr>
<td>In-Service Professional Development</td>
<td>15</td>
</tr>
<tr>
<td>Funding Formula</td>
<td>12</td>
</tr>
<tr>
<td>Other</td>
<td>9</td>
</tr>
<tr>
<td>Teacher Preparation</td>
<td>9</td>
</tr>
<tr>
<td>Teacher Recruitment and Retention</td>
<td>7</td>
</tr>
<tr>
<td>Facilities</td>
<td>7</td>
</tr>
<tr>
<td>Decision Making Authority</td>
<td>4</td>
</tr>
<tr>
<td>Instructional Time/School Hours</td>
<td>4</td>
</tr>
<tr>
<td>Equipment</td>
<td>4</td>
</tr>
<tr>
<td>Class Size/Ratios</td>
<td>2</td>
</tr>
<tr>
<td>School Supplies</td>
<td>2</td>
</tr>
<tr>
<td>School Feeding/Meals</td>
<td>1</td>
</tr>
</tbody>
</table>

- Most common: improve the quality of the teacher and teaching materials
- “Funding” incl.: intervention programs in low performing, low SES schools, financial incentives for teachers and schools and changes in funding allocation (government/private)
- “Other” incl.: School health programs; standards for inputs; transportation; encourage private sector/gov schools partnerships
Teaching/Learning Practices

- Other: 12
- Student-Oriented Pedagogy: 6
- In-Class Learning Strategies: 5
- Tracking/Streaming Policy: 2
- Motivation and Future Plans: 2
- Additional Classes: 2
- Organization of Instructional/Study-Time: 2
- Staff-Collaboration/Mentoring: 2
- Enhanced Learning Activities: 2

Specific approaches championed based on assessment findings include differentiated learning, student oriented pedagogy and in-class peer collaboration.

‘Other’ category covers a broad range of practices, including in-class assessments, mandated use of ICT, checklists of identified best practice, esp. regarding “difficult” items.
Soundness of Program

• The quality of assessment program is the main potential facilitator or barrier to usefulness in policy
  – A third of all studies about *national assessments* considered poor program quality to be a barrier to use
  – On the other hand, a similar proportion find quality of *regional and international assessments* to be facilitators to use, but mainly in influencing (and improving quality of) national assessments
Aspects of Program Quality

• Comparability over time
• Criterion-referenced
• Ability to link/attribute results to student, school and contextual factors that are nationally-relevant
  – Appropriate indicators
  – Taking into account policy goals
  – Sampling to account for sub-national disaggregation of interest
Engaging Stakeholders

• A variety of stakeholders (e.g. NGOs, education lobbyists, conferences), from design (setting priorities) to dissemination; sense of ownership

• Engaging policymakers at national and sub-national levels, especially when education policymaking is decentralised
  – Regional seminars
  – Face-to-face meetings with key policymakers
Other factors influencing use

• Disseminating results to schools that draws out lessons, materials for teachers and schools (e.g. assessment frameworks for classroom use)

• Clearly defined curricular frameworks and standards with potential for relevant comparisons to be made;

• Public opinion often catalyst (e.g. Annual Survey of Education Report ASER, Madhya Pradesh, India; to move from focus on inputs to focus on student outcomes) BUT: Danger of “quick and ineffective policies”
Other factors influencing use

- Integration into the policy process; address policy concerns; meaningful comparisons
- Staggered dissemination to maintain levels of discussion
- Publicly available and well documented databases.
Information Gaps

• (Comparatively) Little is known about:
  – The use/impact of sub-national assessment programs;
  – The use/impact of assessment programs in the Asia-Pacific region;
  – The role of factors outside of the nature of the assessment itself (e.g. characteristics of the educational or political system) in facilitating or impeding assessment data use in policy-making.
Final thoughts

- Desire to examine policy impact of national and international assessment is understandable given their widespread use;
- Rigorous policy analysis scarce;
- Difficult to ascertain policy impact given the link of policy to politics;
- Systematic review using framework analysis: Very involved process.
Sources


The comments from all reviewers of the systematic review protocol are herewith gratefully acknowledged.