Reporting and Dissemination of Large-Scale Learning Assessments: The Bigger Picture

Sally Robertson and Ursula Schwantner
ACER, Centre for Global Education Monitoring (GEM)
Overview

• Introduction to ACER and the Centre for Global Education Monitoring (GEM)
• Elements of a comprehensive assessment system
• Purposes of large-scale assessments and reporting and dissemination
• Relationship of reporting and dissemination to other components of an assessment program
• Workshop topics and discussion
ACER

- Australian Council for Educational Research
- Established in 1930 as a not-for-profit research organisation
- *ACER’s mission*: Creating and promoting research-based *knowledge, products and services* to improve learning
ACER research

Centre for Global Education Monitoring

Centre for Education Policy and Practice

Centre for Assessment Reform and Innovation

Centre for Science of Learning @ ACER

https://www.acer.edu.au/research
ACER international context

- OECD-PISA
- IEA-ICCS
- IEA-TEDS
- International projects in collaboration with Development Partners
ACER Centre for Global Education Monitoring

• The GEM Centre supports the monitoring of educational outcomes worldwide to inform evidence-based education policy aimed at improving educational progress for all learners.

• Collaboration between ACER and DFAT
GEM key areas

System strengthening
Learning metrics
Reviews and analysis
Monitoring Trends in Educational Growth partnership service

• MTEG is a collaborative approach to develop and implement an assessment that meets the specific monitoring needs of a country.

• MTEG is based on a set of defined design principles and quality standards, yielding high-quality assessment data of national relevance.

https://www.acer.edu.au/gem/key-areas/system-strengthening/mteg
Assessment purpose

“The fundamental purpose of assessment in education is to establish and understand where learners are in an aspect of their learning at the time of assessment.” (Masters, 2014, p. 1)
Components of an assessment system

- Classroom based assessments
- School based assessments
- National based census assessments
- National based survey assessments
- International and regional assessments
Classroom based assessments

• For monitoring the progress of individuals and providing information in determining appropriate pedagogical response
School based assessments

- For monitoring the progress of individuals or to aid the selection of future educational pathways
- Could be single school assessments or district based to allow comparisons of progress
National based census assessments

- For monitoring the progress of individuals, for accreditation purposes, or to aid the selection of future educational pathways
- Can be used for sub-population monitoring purposes
National based survey assessments

• For monitoring the health of the education system
• For making sub-national comparisons of educational outcomes such as between provinces
International and regional assessments

- For benchmarking national performance against that of other countries
- Can be used for sub-population monitoring purposes
International assessments
Regional assessments

Source: http://www.seaplm.org/seaplm/index.php/countries
Participation in international, regional and national assessments

- IEA conducted first internationally comparative study in mathematics with 12 countries in 1964
- There has been a growth in participation in all large-scale assessments, particularly amongst developing countries (Kamens and McNeely, 2009)
- Between 2007-2013, 67% of countries in the Asia-Pacific region conducted a national assessment (Benavot and Köseleci, 2015)
Large-scale assessments

• The focus of the workshop this week is on the reporting and dissemination of large-scale assessments.

• Large-scale assessments are:
  – Representative of an education system
  – Standardised
  – Not national examinations
  – National, regional and international assessments

<table>
<thead>
<tr>
<th></th>
<th>Examination</th>
<th>Monitoring Assessment</th>
</tr>
</thead>
</table>
| **Purpose**                         | Ranking students to select the most able students for the next stages of education and/or to certify students  
Measuring the degree to which curricular content has been learned | Quantify differences in learning between sub-populations  
Measure changes in learning over time  
Establish and describe the range of what students know |
| **Type**                            | Census                                                                                           | Sample or census                                                                        |
| **Content**                         | Based entirely on the intended curriculum of a single grade level                               | May be inside or outside of the curriculum  
May be at higher or lower levels of the intended curriculum for a particular grade |
| **Approach**                        | Focus on whether the skills have been attained and the content has been learned                  | Focus on whether the learner can apply the skills and concepts that she or he has acquired and learned |
| **Analysis**                        | Simple, percent-correct scores                                                                  | Complex statistical methods are required                                                 |
| **Reporting**                       | Reports at the individual level                                                                 | Reports at the level of populations and sub-populations                                  |
Evaluating and monitoring the quality of the education system

Ensuring equity

Exercising accountability

Policy impacts of large-scale assessments

Factors that influence the utilisation in education policy

Integration into policy processes

Media and public opinion

Quality of the assessment program

Key areas of a robust assessment

Source: ACER-GEM (https://www.acer.edu.au/gem/about/approach)
Policy goals and issues

• Large-scale assessments should be designed with the aim of obtaining data to help address policy issues of interest and intended policy goals.
Project team and infrastructure

• Essential to the success of any large-scale assessment are:
  – Adequate number of skilled staff to cover the different roles needed (e.g., project management, item development, sampling, analysis, reporting and communications)
  – Adequate resources and a core physical infrastructure
Technical standards

The technical quality of the assessment can act as a barrier or facilitator to the use of assessment results in education decision-making (Tobin et al., 2015).
Technical standards

• Guidelines to ensure the consistency, precision and generalisability of the data, for example:
  – sampling standards
  – data standards
  – psychometric standards

• The overall goal is to create a dataset of a quality that allows valid inferences to be made.
Assessment framework

- Explicit statement and discussion about what an assessment intends to measure
- Lays out the principles upon which an assessment is built
- Guides test development

Contextual framework

• Based on theory or evidence from research literature
• Formulates and addresses important policy questions and issues
• Guides the questionnaire development
• Guides the analysis and reporting
### Two-dimensional taxonomy of educational outcomes and predictive factors

<table>
<thead>
<tr>
<th></th>
<th>Input</th>
<th>Processes</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Students</strong></td>
<td>Gender, grade level, socio-economic status</td>
<td>Attendance/truancy</td>
<td>Mathematical performance</td>
</tr>
<tr>
<td></td>
<td>Educational career, grades</td>
<td>Outside-class activities - e.g. participation in after-school programmes</td>
<td>Mathematics-related attitudes, beliefs and motivation</td>
</tr>
<tr>
<td></td>
<td>Immigration background, family environment and support</td>
<td>Motivation, engagement</td>
<td>General school-related attitudes and behaviour, e.g. commitment, truancy</td>
</tr>
<tr>
<td></td>
<td>ICT experience, attitudes, skills</td>
<td>Learning and thinking strategies, test taking strategies</td>
<td>Learning motivation, educational expectations</td>
</tr>
<tr>
<td></td>
<td>Openness, problem-solving styles</td>
<td>Learning time (including homework and private tuition)</td>
<td></td>
</tr>
<tr>
<td><strong>Classrooms</strong></td>
<td>Class size, socio-economic background and ethnic composition</td>
<td>Quality of instruction: structure, support, challenge</td>
<td>Aggregated student variables</td>
</tr>
<tr>
<td></td>
<td>Teacher education/training, expertise</td>
<td>Opportunity to learn: implemented curriculum, assigned tasks, mathematics-related activities</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Instructional time, grouping, assessment and feedback</td>
<td></td>
</tr>
<tr>
<td><strong>Schools</strong></td>
<td>Socio-economic background and ethnic composition</td>
<td>Achievement orientation, shared norms, leadership, teacher morale and co-operation, professional development</td>
<td>Aggregated student variables</td>
</tr>
<tr>
<td></td>
<td>Affluence of the community</td>
<td>Admission and recruitment policies, tracking, course offerings/school curriculum, evaluation</td>
<td>Promotion/retention and graduation rates</td>
</tr>
<tr>
<td></td>
<td>School funding, public vs. private</td>
<td>Teacher-student relations, supportive environment</td>
<td>Attendance</td>
</tr>
<tr>
<td></td>
<td>School size</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parental involvement</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Countries</strong></td>
<td>Economic wealth, social (in)equality</td>
<td>School funding, tracking and allocation, policies for professional teacher development, support for special needs and language minority students, hiring and certification policies</td>
<td></td>
</tr>
<tr>
<td><strong>(Systems)</strong></td>
<td>Diversity policies</td>
<td>Accountability and evaluation policies, locus of decision making</td>
<td>Aggregated student variables</td>
</tr>
</tbody>
</table>

OECD, 2013, p. 175
High quality assessment items...

• meet the requirements of the assessment framework
• cover a range of skills
• are well targeted to students’ abilities
• are fair
• will improve the explanatory power of the assessment and increase usefulness of the assessment findings
Linguistic quality control

- Where assessments are administered in different languages, the different versions should be equivalent.
Test design

Exhibit 1: MTEG assessment schedule in Afghanistan

Sample design

- To enable valid inference beyond the students assessed to a population of students, scientific sampling techniques must be employed.
Standardised field operations

- Standardised test administration procedures are vital to ensure that every student is assessed under the same conditions.
Data management

- Quality control processes should be implemented at all stages of data management including:
  - data entry and processing
  - creating the database
  - preparing the data for reporting

- It is essential to provide full documentation with the database.
Scaling methodology

Lumley, Mendlovits, Stanyon, Turner & Walker (2015)
Data analysis

• Analysing data from a complex survey design requires advanced statistical techniques.
• Data analysts need to be highly skilled to appropriately apply and interpret the analyses.
• A data analysis plan should be developed to inform the assessment design and to provide information about the kind of findings that will be possible.
Reporting and dissemination

• Reporting and dissemination should be guided by an effective strategy, and incorporate a range of approaches to cater for a diversity of audiences.

• The main aim of all reporting should be to support an understanding of the assessment and its findings, in response to the policy goals defined for the assessment program.
Reporting and dissemination in the assessment cycle

Source: ACER-GEM (https://www.acer.edu.au/gem/about/approach)
What is happening over the next three days?

**Wednesday**
- Dissemination approaches for different purposes and audiences
- Developing a dissemination strategy
- Developing assessment products

**Thursday**
- Reporting achievement and contextual data
- Interpreting assessment results
- Translating assessment findings into policy and practice

**Friday**
- Enhancing public awareness of assessment results
- Assessment websites as a dissemination tool
Discussion

• What is your role in assessment reporting and dissemination?

• What dissemination products have you previously been involved in developing/reviewing/editing?

• What challenges have you faced with reporting and dissemination? How did you address these challenges?

• What is one question you’re hoping to have answered during this workshop?


For information about ACER see: https://www.acer.edu.au/

For information about ACER’s GEM Centre see: https://www.acer.edu.au/gem