Incorporating out-of-school youth in international large-scale assessments – the case of PISA for Development

Michael Ward
OECD Development Co-operation Directorate & Directorate for Education and Skills

World Bank, Washington – 3-4 February 2020
Measuring Learning in National Household Surveys
This presentation

• PISA and PISA for Development, a brief introduction
• Rationale for a PISA household survey
• The assessment structure (sampling, questionnaires, skills and subjects covered)
• Data collection design
• Results
• Technical/operational challenges
• Lessons learned and next steps
Key documents
The Programme for International Student Assessment (PISA)

- **15-year-old students in 7th grade** or above from **randomly selected schools** take two-hour tests (CBA and PBA) in the key subjects: **reading, mathematics and science** plus optional assessments.

- Tests are **competency based and internationally comparable**.

- **Background questionnaires** for students, schools, teachers and parents provide context which can help interpret the results.
PISA design

- Launched in 1997 by OECD - assessments conducted every three years since 2000 - Reading, Mathematics, Science and other optional domains
- Major/minor domains

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Read</td>
<td>Read</td>
<td>Read</td>
<td><strong>Read</strong></td>
<td>Read</td>
<td>Read</td>
<td><strong>Read</strong></td>
<td>Read</td>
<td>Read</td>
</tr>
<tr>
<td>Math</td>
<td><strong>Math</strong></td>
<td>Math</td>
<td>Math</td>
<td><strong>Math</strong></td>
<td>Math</td>
<td>Math</td>
<td><strong>Math</strong></td>
<td></td>
</tr>
<tr>
<td>Science</td>
<td>Science</td>
<td><strong>Science</strong></td>
<td>Science</td>
<td>Science</td>
<td><strong>Science</strong></td>
<td>Science</td>
<td>Science</td>
<td></td>
</tr>
<tr>
<td>Prb Slv</td>
<td>Prb Slv</td>
<td>Prb Slv</td>
<td><strong>Collab. Prb. Slv</strong></td>
<td>Global Competence</td>
<td>Creative thinking</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Participating countries and economies in PISA 2018

Over half a million 15-year-olds from 79 countries and economies

All 37 OECD member states and 42 partner countries/economies

OECD countries
Partner countries and economies
EXAMPLES OF PISA OUTCOMES – THREE MAIN TYPES
Basic indicators: Mean reading performance

Tab I.4.1

Countries/economies statistically significantly above the OECD average
- Macao (China)
- Estonia
- Ireland
- Poland
- Japan
- Germany
- France
- New Zealand
- United States*
- Canada
- Finland
- Korea
- Sweden
- Norway
- Chinese Taipei
- Singapore

Countries/economies statistically significantly different below the OECD average
- Hong Kong (China)*
- Australia
- Chinese Taipei
- Macao (China)
- Brazil
- Colombia
- Costa Rica
- Mexico
- Montenegro
- Montenegro
- Peru
- Argentina
- Albania
- Qatar
- North Macedonia
- Kazakhstan
- Panama
- Kosovo
- Morocco
- Lebanon
- Dominican Republic

Countries/economies NOT statistically significantly different from the OECD average
- United Kingdom
- Slovenia
- Belgium
- Portugal*
- Netherlands*
- Croatia
- Latvia
- Hungary
- Italy
- Israel
- Ukraine
- Slovak Republic
- Chile
- Malta
- Serbia
- United Arab Emirates
- Uruguay
- Moldova
- Montenegro
- Jordan
- Malaysia
- Brazil
- Brunei Darussalam
- Bosnia and Herzegovina
- Saudi Arabia
- Thailand
- Baku (Azerbaijan)
- Georgia
- Indonesia

*Countries/economies with an asterisk* did not meet response-rate standards, but further analyses could exclude a large bias in the published results due to non-response.
Contextual indicators: Reading performance and equity in PISA 2018

Some countries combine equity and excellence

Low performance
Low equity

Greater equity

Higher Performance

High performance
High equity
Indicators on trends: Change between 2015 and 2018 in mean reading performance

Fig I.8.1

Statistically significant differences between PISA 2015 and PISA 2018 are shown in a darker tone
RATIONALE FOR A PISA HOUSEHOLD SURVEY – PISA FOR DEVELOPMENT

PART ONE
Three principle drivers of PISA for Development (1)

- OECD Strategy on Development 2012
- *make OECD instruments available to a wider range of countries*
### Three principle drivers of PISA-D (2)

#### Increasing demand to participate in PISA

<table>
<thead>
<tr>
<th>PISA cycle</th>
<th>Members</th>
<th>Non-Members</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>28</td>
<td>15</td>
<td>43</td>
</tr>
<tr>
<td>2003</td>
<td>30</td>
<td>11</td>
<td>41</td>
</tr>
<tr>
<td>2006</td>
<td>30</td>
<td>27</td>
<td>57</td>
</tr>
<tr>
<td>2009</td>
<td>34</td>
<td>40</td>
<td>74</td>
</tr>
<tr>
<td>2012</td>
<td>34</td>
<td>30</td>
<td>64</td>
</tr>
<tr>
<td>2015</td>
<td>35</td>
<td>36</td>
<td>71</td>
</tr>
<tr>
<td>2018</td>
<td>37</td>
<td>42</td>
<td>79</td>
</tr>
<tr>
<td>2021</td>
<td>35</td>
<td>50</td>
<td>85</td>
</tr>
</tbody>
</table>
Three principle drivers of PISA-D (3)

- Inclusion of PISA as a source of data for global monitoring of SDG 4.

Global Indicator 4.1.1.c

Proportion of children and young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex

- Equates to: Level 2 in PISA (at least 407 points for reading; 420 points for mathematics)
PISA-D informed by the experience of the many low- and middle-income countries already in PISA...

- PISA results highlight large differences in student performance
- Some of the contextual factors measured by PISA are not adequately capturing the social and economic context in which students learn, teachers work and schools operate in lower and middle-income countries
- Because out-of-school rates are high in middle income countries indices of coverage in these contexts can be as low as 50%; in low income countries average indices of coverage are 30%
PISA-D Implementation: 2014-2020

- Nine countries from Africa, Asia and Latin America
- OECD and international contractors
- Development partners
- In-school and out-of-school components
- Capacity building of participating countries
- International and national reports
PISA is now more relevant to Low and Middle Income Countries

- Increased the resolution of the PISA tests at the lower end of the student performance distribution – Level 2 and below
- Captured a wider range of social and economic contexts
- Incorporated an assessment of out-of-school 14-16-year-olds

PISA now more accessible to Low and Middle Income Countries

- Partnerships for participation established
- Better preparation of countries and their populations
- Capacity building and peer-to-peer learning established
- Collaboration regarding analysis of data, reporting and dissemination and use of results established
What changes in PISA after PISA-D?

PISA-D informed by:
• PISA experience
• Participating countries
• Other assessments

Outputs of PISA-D (instruments & approach for incorporating out-of-school youth in PISA) enhancing PISA from 2021 cycle onwards
PISA-D a success, but the results reveal the extent of the global learning crisis

(PISA-D, reading, 15-year-olds students in Grade 7 or above)

<table>
<thead>
<tr>
<th>Country</th>
<th>Percentage</th>
<th>Level 1a</th>
<th>Level 1b</th>
<th>Level 1c</th>
<th>Below Level 1c</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
<th>Level 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>OECD average</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zambia</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Senegal</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paraguay</td>
<td>m</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Honduras</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guatemala</td>
<td>53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ecuador</td>
<td>39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cambodia</td>
<td>72</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Percentage of 15-year-olds not included in the PISA sample.
RATIONALE FOR A PISA HOUSEHOLD SURVEY

PART TWO
What was the problem we set out to address? Percentage of 15-year-olds covered by PISA

On OECD average, PISA 2018 represents 88.2% of the entire 15-year-old population across OECD countries.
Because out-of-school rates are high in many countries, indices of coverage in low-and-middle income countries especially can be as low as 30%, we wanted to achieve:

- An approach and methodology for incorporating out-of-school youth in PISA assessments
Using a sample of in-school eligible students as a proxy for a cohort of 15 year olds or an entire schooling system can lead to misleading results in contexts where large numbers of youth are out-of-school.

Inaccurate cross-country comparisons arise when using only PISA data, especially when countries have different levels of incomplete access or different proportions of delayed (& ineligible) students.

Underestimating progress of the real educational improvements over time for countries that have improved access/attainment.

Underestimating inequality When there is a sample selection process involved such that poorer students are more likely to be excluded from the sample, PISA will underestimate socioeconomic inequalities.

Potential perverse incentives for countries to maintain policies of exclusion in order to inflate performance on international tests
In the absence of a solution...

- estimate test scores for the whole population (i.e., taking into account dropouts) by putting bounds on unobserved scores...
- this is effectively guess-work and is carried out under assumptions that are not underpinned by real evidence...
- continuing potentially perverse incentives for countries to exclude low-performing students from schooling...
- in context of SDG 4 *(leave no one behind)*, there is no substitute for assessing the skills of the whole population.
PISA eligible students

• students aged between 15 years and 3 (completed) months and 16 years and 2 (completed) months at the beginning of the testing period, attending educational institutions located within the adjudicated entity, and in grade 7 or higher.
PISA-D out-of-school target population

- 14-16 year-old youth who are either enrolled in school at grade 6 or below or who are outside of the school system (PISA-D country averages)
PISA-D out-of-school assessment pilot achievements

- Counted and located the target population (sampling frame)
- Found and identified the target population (sampling strategy)
- Developed and implemented an assessment of reading and mathematics delivered in the household on a tablet computer
- Developed and implemented contextual questionnaires delivered in the household
- Administered a survey in the most cost-effective way, given the strategy
- Linked the results to the PISA scale
- Achieved enough completed cases to test the validity of the items and allow analyses that are useful to the pilot and relevant for the countries
- Will report on achievements and lessons learned
Went *much lower* on the reading scale

<table>
<thead>
<tr>
<th>Illustrative examples</th>
<th>Reading</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 1c</strong></td>
<td>• Decodes and understands short sentences (&quot;The red car has a flat tyre&quot;, &quot;airplanes are made of dogs&quot;)</td>
</tr>
<tr>
<td><strong>Level 1b</strong></td>
<td>• Understands short text, finds a single piece of explicitly stated information (e.g. &quot;what colour is the car?&quot;)</td>
</tr>
<tr>
<td><strong>Level 1a</strong></td>
<td>• Level 1b + Identifies the main theme or the author’s intent in a text about a familiar topic</td>
</tr>
<tr>
<td><strong>Level 2 (baseline)</strong></td>
<td>• Reads and understands simple texts; • connects pieces of information, draws inferences beyond the explicitly stated</td>
</tr>
</tbody>
</table>
Went *much lower* on the mathematics scale

<table>
<thead>
<tr>
<th>Illustrative examples</th>
<th>Mathematics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level 1c</strong></td>
<td>• What is the price of orange juice at this restaurant?</td>
</tr>
<tr>
<td><strong>Level 1b</strong></td>
<td>• Which drink is most expensive?</td>
</tr>
<tr>
<td><strong>Level 1a</strong></td>
<td>• How much do you pay if you order 2 orange juices and a snack?</td>
</tr>
<tr>
<td><strong>Level 2 (baseline)</strong></td>
<td>• How much cheaper is the « breakfast deal » compared to ordering each item separately from the menu?</td>
</tr>
</tbody>
</table>
Enhanced the background questionnaires

• In addition to the PISA-D student questionnaire elements, obtained information about why the youth is not in school, barriers preventing the youth from returning to school, and about employment.

• Through the Parent (or the most knowledgeable person) questionnaire, obtained more information about the youth’s background and childhood experiences.

• Expanded measure of economic, social and cultural status (ESCS) to adequately capture lower levels of parental education, income and risk factors of poverty
The assessment structure (1)

Results from out-of-school linked to the scales used in in-school, thus requiring a large proportion of overlapping items between the two surveys.
The assessment structure (2)

• Administration of out-of-school assessment via tablets
• Maximized the use of automatically scored items to capitalize on the use of tablets
• Focused on reading and mathematics only
• A routed design with two paths: a cognitive path more similar to PISA assessments of in-school populations, or to path with a set of tasks resembling components
• Youth interviewed first for completion of background questionnaire and then takes the test
Data Collection Design

**Respondent**
- **In-person interview**
  - **Youth Interview**
    - (30-35 Min)
  - **Core Module**
    - (5 Reading and 5 Mathematics Items)
      - (10 Min)
  - **Reading Components**
    - (Sentence Processing and Passage Comprehension)
      - (15 Min)
  - **Forms 1-12**
    - (Combination of Reading Components, Reading, and Mathematical Literacy Items)
      - (35 Min)

**Others**
- **Person(s) most knowledgeable about the respondent questionnaire (i.e., parents, caregivers, Guardians)**
- **Household Observation Schedule (Interviewer)**
Sampling and survey operations

Sampling approaches used

1. Representative Probability

2. Link-tracing through householder referrals or recruiting

3. Limited Representative
   a) School frame approach for out-of-school youth
   b) School frame approach for <7th grade
   c) Location sampling

Survey operations

- National Centres in each country
- Liaison with National Statistics Offices
- Field Trial and Main Survey
- 60-100 interviewers per country
Results Snapshot - the final reports will be published in April 2020

- There is no hidden wealth of literacy and numeracy in the out-of-school youth population
- The difference between the performance on the test of the in-school compared to the out-of-school is the equivalent of between 1-2 years of schooling
- The performance on the test of out-of-school youth in formal employment compares favourably to the in-school
- Almost all of the out-of-school youth that had dropped out or were in-school but in grade 6 or below had repeated early primary grades
Technical/operational challenges faced (data analysis is on-going)

• This target population has unique characteristics and low levels of proficiency that must be considered during future designs

• The adequacy of the PISA frameworks for this unique population should be examined further

• The complexities associated with identifying these populations and assessing them through household surveys present unique challenges

• The timing of the out-of-school assessment in the context of a PISA cycle should be examined further

• The effort required from National Centers for the implementation of this survey should be recognized and participating countries applauded
Key lesson learned

• The approach and methodology works, but ...

• out-of-school assessment in households is expensive and main in-country costs are those related to identifying and locating respondents...

• ...a large amount of screening required to locate eligible youth and good local area data is essential...

• ...these costs of screening are prohibitive and will constrain scaling-up of the initiative unless solutions are found to screening challenge..
PISA-D out-of-school assessment next steps

- Country reports launching in **Panama** (December 2019) and **Guatemala, Honduras, Paraguay** and **Senegal** in March/April 2020

- Technical Report, International Dataset and PISA-D Results-in-Focus launched by OECD in March/April 2020

- *Technical workshop at OECD, Paris from 3-5 June 2020* to review lessons learned from pilot and to agree on improvements and plans for scaling up the assessment in PISA 2024 and beyond – stand-alone survey or integrated with household surveys

- Explore with countries, household survey managers and partners a *road-map for PISA out-of-school assessment* to potentially *piggyback* on household surveys to reduce the cost of screening
Potential countries for out-of-school assessments: PISA 2018 participants with coverage index 3 values of 70% or less

- Brazil
- Baku (Azerbaijan)
- Colombia
- Costa Rica
- Jordan
- Mexico
- Morocco
- Panama (has already piloted an out-of-school assessment in PISA-D)
- Philippines
- Viet Nam